8-1-2011

Predictors of Performance In A Professional Counselor Masters Program

Mya M. Bethune

UNIVERSITY OF NORTHERN COLORADO

Greeley, Colorado

The Graduate School

PREDICTORS OF PERFORMANCE IN A PROFESSIONAL COUNSELOR MASTERS PROGRAM

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

Mya M. Bethune

College of Education and Behavioral Sciences
School of Applied Psychology and Counselor Education
Counseling Psychology Program

August, 2011

This Dissertation by: Mya M. Bethune

Entitled: *Predictors of Performance in a Professional Counselor Masters Program*

has been approved as meeting the requirement for the Degree of Doctor of Philosophy in College of Education and Behavioral Sciences in School of Applied Psychology and Counselor Education, Program of Counseling Psychology

Accepted by the Doctoral Committee

Brian Johnson, Ph.D., Chair

David Gonzalez, Ph.D., Committee Member

Stephen Wright, Ph.D., Committee Member

Daniel Mundfrom, Ph.D., Faculty Representative

Date of Dissertation Defense

Accepted by the Graduate School

Robbyn R. Wacker, Ph.D
Assistant Vice President for Research
Dean of the Graduate School

ABSTRACT

Bethune, Mya. *Predictors of Performance in a Professional Counseling Masters Program*. Published Doctor of Philosophy dissertation, University of Northern Colorado, 2011.

The prediction of success in graduate education is of growing interest in the field of professional psychology. Large scale studies and meta-analyses have established that cognitive and personality variables do provide increments of validity in the prediction of individual differences in academic outcomes. However, there continues to be debate regarding the most effective variables to build predictive models and how to define success in graduate school. This dissertation analyzed whether the prediction of performance in a professional counseling masters program (PCMP) using undergraduate grade point average and Pre-Admissions Workshop rating scores could be incremented by adding personality traits, as measured by the MMPI-2. Graduate grade point average (GGPA) and internship evaluation ratings (IER) were both used as success criteria in order to investigate whether a stronger predictive model could be built using a traditional outcome variable or a criterion more in line with the training objective of a professional program. The two hierarchical regression analyses produced a number of key findings. Across both analyses, the MMPI-2 independently accounted for the most variance in performance outcome, after controlling for the traditional admissions variables. The final model accounted for 24.4% of the variance in GGPA and of that, the MMPI-2 uniquely explained 17.7%. When IER was used as the criterion, the final model accounted for

15.6% of the variance, and 14.5% of that variance was uniquely explained by the MMPI-2. Scales 9 (Hypomania), 2 (Depression) and F (Infrequency) on the MMPI-2 were the most significant contributors to the prediction of performance. Overall, these findings provided empirical support for the inclusion of an objective, valid personality instrument in PCMP admissions procedures and suggest that personality characteristics are veritably relevant to academic performance. Likewise, the use of IER as a criterion of success in a PCMP demonstrated potential as a variable that could overcome the limitations of using GGPA. Suggestions as to how to build on these results, through future research, are provided.

ACKNOWLEDGEMENTS

As I complete this dissertation and reflect on the long journey pursuing this doctorate degree, there are many people who come to mind that I would like to express my immense appreciation for their incredible kindness, understanding, guidance and support.

I would thank my academic advisor and the Chair of my doctoral committee, Dr. Brian Johnson, for his supportive advisement, words of encouragement, thoughtful feedback and time during incredibly busy semesters. Likewise, I appreciate his instruction in psychological assessment as it was paramount in inspiring this dissertation topic and my passion for that type of work.

I would also like to thank Dr. David Gonzalez, for supporting my professional development even after retirement. Having him as a committee member and a professor has been an honor and a privilege. Not only have I benefitted from his mentorship, but I am also certain that my clients already have and will continue to benefit from the positive influence he has had on my work as a clinician.

Additionally, I extend my gratitude to Dr. Daniel Mundfrom, for his input, patience and accessibility. I truly appreciated his willingness to translate statistical concepts into terms I could comprehend and especially for his time consulting on this research project. I would also like to thank, Dr. Steven Wright, for taking a chance by agreeing to be on my committee when he first began his career at UNC. I appreciate all the support and feedback he provided during such busy semesters.

Widening the scope, I am beholden to my wonderful extended family and friends, larger than I can thank by name. I am so grateful for the understanding, humor and support you have all provided along the way. Notably, my closest friend, Kristalyn Larsen, whose presence in my life is a gift.

Lastly, my deepest gratitude goes to my amazing family. To my husband, Alec, who has always been by my side, providing unwavering encouragement and strength without me ever having to ask. I am so grateful for his unconditional love and support that has made this journey possible. Kate and Julia, my beautiful daughters, whose exuberance towards life, kind nature and joyful spirit continuously inspire me to be a better person and to always strive to be my best. Lastly, I extend my appreciation towards Bentley, Callebaut and Max, our beloved dogs, for keeping me company as I often worked late into the night. There is nothing more important to me than our family and I am so unbelievably fortunate to have all of you in my life.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Over the past twenty years, there has been considerable focus on competency in mental health literature. This is especially evident in relation to screening trainees for admissions and assessing individuals once they have been accepted into professional psychology graduate programs. The definition of competence utilized by the American Psychological Association's Assessment of Competency Benchmarks Work Group (2007) is as follows, "professional competence is the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values and reflection in daily practice for the benefit of the individual and community being served" (Epstein & Hundert, 2002, p.226). To be a competent psychology graduate student, therefore, requires certain coalescence in the dimensions of ability, aptitude and application. As is the case with all professions, the definition and measurement of those constructs is both philosophically complex and debatable.

Empirical research in the field of applied psychology on predictor – criterion relationships and incremental validity provides methods with which to better understand constructs and the relationships between constructs. In fact, research focused on the relationship between personality as well as intelligence and academic achievement has been forthcoming since the early 1900's (Poropat, 2009). During World War I, the scope of research on predictor – criterion relationships widened as psychologists in the United

States began using objective personality assessments for personnel selection (Butcher, Ones & Cullen, 2006). Rouse and Butcher (1995) reviewed the research supporting the use of the Minnesota Multiphasic Personality Inventory (MMPI) and MMPI-2 in personnel settings and found over 570 articles and books. Accordingly, Butcher et al. (2006) asserted, "The use of psychological tests in making personnel –related decisions is an appropriate professional activity for psychologists to undertake and a valuable contribution to the employment process" (p.381).

When using a personality assessment for screening, psychologists typically interpret findings based on inclusion and exclusion rules. Does the candidate have characteristics that match desired attributes for a position or possess a personality profile similar to those with significant mental health concerns? The importance of this type of screening is underscored when it comes to positions of significant public trust such as airline pilots, police officers, firefighting personnel, medical and clinical psychology students (Butcher, 1979). Although there is an abundance of empirical research to support the use of the MMPI and MMPI-2 for personnel screening, it is not standard practice to administer an objective personality assessment as part of admissions to professional psychology graduate programs. Rather, these characteristics tend to be qualitatively assessed with methods such as letters of recommendation, personal interviews and personal statements.

There are no set standards in terms of admissions criteria for graduate programs in professional psychology; however, the most common and heavily weighted criteria are undergraduate grade point average, letters of recommendation and personal statements (Norcross, Kohout & Wicherski, 2005). Johnson and Cambell (2004) reported that

directors of clinical training programs endorsed being significantly concerned about both the moral character and psychological fitness of trainees. However it is echoed throughout the literature and specifically stated by Johnson and Cambell that the reason personality characteristics are not objectively assessed is because there is a lack of research indicating that it can validly predict performance in professional psychology graduate programs.

Additional support for the objective assessment of personality factors as part of admissions comes from research validating that as training in rigorous graduate programs progresses, the mental health or psychological fitness of trainees becomes a more salient predictor of success (Lievens, Dilchert & Ones, 2009). This seems to be related to the translation of knowledge into applied practice and interpersonal skills related to working with people. Similarly, the literature indicates that individuals who demonstrate difficulty with behaviors of competency in professional counseling masters programs struggle in the application of knowledge (e.g., clinical skills) rather than in the acquisition of declarative or procedural knowledge (Gaubatz & Vera, 2002; Olkin & Gaughen, 1991; Vacha-Haase, Davenport & Kerewsky, 2004). Furthermore, Kuyken, Peters, Power and Lavender (2003), found that as individuals progress in clinical psychology programs, the overall number of distressed students increase and the level of severity of their distress increases as well.

Statement of the Problem

There is consensus in the literature that intelligence is a strong predictor of academic success and that the predictive validity of future academic performance decreases as individuals progress to higher levels of education (Busato, Prins, Elshout &

Hamaker, 2000; Di Fabio & Palazzeschi, 2009; Farsides & Woodfield, 2003; Furnham & Monsen, 2009; Trapmann, Hell, Hirn & Schuler, 2007). This is especially the case with graduate school; however, there is a lack of predictive validity research at this level of education. Large scale studies as well as meta-analyses have evidenced that the prediction of academic performance using cognitive variables (e.g., GPA or academic achievement tests) can be incremented by adding objective measures of personality (Lievens, Dilchert & Ones, 2009; Poropat, 2009; Trapmann et al., 2007). In fact, Poropat (2009), who conducted the most comprehensive meta-analytic investigation of personality-academic performance relationships concluded, "Personality should take a more prominent place in future theories of academic performance and not merely as an adjunct to intelligence" (p.333).

The present study attempted to contribute to this line of inquiry by diversifying the analysis. Whereas the existing research utilized the Five Factor Model (FFM) to conceptualize personality traits, the present study used the MMPI-2. Likewise, previous studies examining the predictive validity of personality traits and academic success were conducted for research purposes and after students were admitted. The significance of this study; therefore, is that it investigated how well the MMPI-2 predicts performance when it had been implemented as part of admissions screening procedures in a professional counseling masters program.

Another well documented issue in predictor- criterion and incremental validity research is that there is a lack of consensus as to what constitutes "success" in graduate professional programs (Goldberg & Alliger, 1992; Trapmann et al., 2007). As previously delineated, the definition for competency utilized by the APA highlights that success in a

graduate psychology program is not unidimensional. Despite the multi-faceted skill set necessary for competence and subsequent "success," the most common criterion used to define success is grade point average (GPA) (Kuncel, Crede & Thomas, 2005). GPA has been empirically supported as a useful measure of academic performance as studies in secondary and tertiary education demonstrate it remains reliable over time and has criterion validity (Poropat, 2009). Nonetheless, existing literature indicated the need to identify other criteria of college success as there are notable statistical concerns when using GPA such as range restriction, grade inflation, and the ability to account for attrition. Likewise, GPA data will not reflect variance in performances in practicum or internship if those courses do not factor into the GPA score numerically (i.e., pass/fail).

Purpose of the Study

The purpose of this study was twofold. Given that there is a lack of consensus as to whether personality characteristics should be objectively assessed as part of standard admissions to graduate school, the first objective of the present study was to provide empirical evidence to support the inclusion of valid, reliable personality instruments in professional counseling masters programs (PCMP) admissions procedures. This was accomplished by testing the hypothesis that the MMPI-2 could predict final grade point average and provide incremental validity above what could be explained by previous academic achievement; namely undergraduate grade point average. The multiple regression analysis also included ratings given by the admissions committee members during the Pre-Admissions Workshop in order to test whether adding an objective personality instrument could strengthen the predictive validity of standard admissions criteria, or if a qualitative rating was a stronger predictor of future success in the PCMP.

Testing that hypothesis contributed to the existing literature in several significant ways. First, it met the need to identify other predictive variables of success in graduate school where range restriction is perhaps most prevalent. Second, this study diversified the research as the sample was administered the MMPI-2 as part of admissions procedures. This is notable given the analysis was able to test whether that screening procedure contributed to the selection of successful candidates and whether there were salient personality characteristics that were overlooked as indicators of future performance. Third, there is an abundance of large scale studies and meta-analyses that have established the predictive and incremental validity of personality characteristics and academic performance; however, all of this research used instruments based on the Five Factor Model (FFM). The present study, therefore, contributed to that line of inquiry by testing whether the MMPI-2 could add incremental validity as did instruments based on the FFM and whether the MMPI-2 could explain even more of the variance in academic performance. Fourth, the analysis provided important information regarding the utility of including the MMPI-2 as part of PCMP admissions procedures because the amount of variance explained in performance was directly comparable to standard Pre-Admissions Workshop ratings. In other words, was there value in quantitatively assessing personality characteristics or was as much variance explained by the more common qualitative assessments made by admissions committee members?

Further significance of this analysis was that it integrated recommendations from prior research by using a sample that compared students in a single academic discipline, with a standardized curriculum with which to compare grade point averages. In other words, course difficulty and grading standards should have been more consistent and

reliable than would have been the case if different academic disciplines or programs from other universities were part of the sample.

The second goal of the present study was to meet the need cited in the literature to test criteria other than GPA as a measure of success and overcome the limitations of using GPA as the criterion in a graduate program where application of knowledge becomes a quintessential indicator of success. This was especially significant for this sample where the variance in performance on practicum and internship was not reflected in GPA scores because it was factored in dichotomously as either satisfactory or unsatisfactory. For example, one might have to repeat a practicum or struggle in an area of internship (e.g., openness to feedback) but that individual's struggle to demonstrate competence was not reflected numerically in GPA. That variance in performance was captured on internship evaluations, which is one of the reasons a separate regression analysis was conducted using the same predictor variables and scores on internship as the criterion.

Additionally, internship evaluation scores were analyzed because the form is standardized and the literature suggests that it is a more accurate reflection of performance in a PCMP. Daehnert and Carter (1987) explained,

> Overall, Internship Evaluations are ideal criteria for graduate study in psychology. Especially valuable is the fact that these evaluations are measuring professional activities as perceived by professionals that are actually practicing in the field of clinical psychology. Equally important is the fact that these professionals are functioning outside the graduate community, which provides evaluations that are uncontaminated by program expectations or by students' past performance (p.1122).

Therefore, the second analysis contributed to the literature and benefited PCMPs in a similar manner as the first. The main purpose; however, was to test the amount of

variance that could be explained in performance when using a criterion more aligned with the training objectives of a PCMP.

Research Hypotheses

H1 The prediction of graduate grade point average in a professional counseling masters program using undergraduate grade point average and Pre-Admissions Workshop rating scores can be incremented by adding personality traits, as measured by the MMPI-2.

H2 The prediction of internship evaluation ratings in a professional counseling masters program using undergraduate grade point average and Pre-Admissions Workshop rating scores can be incremented by adding personality traits, as measured by the MMPI-2.

H3 Personality traits, as measured by the MMPI-2, will have more predictive validity when internship evaluations rather than graduate grade point averages are used as the criterion.

Limitations

There were limitations to the present study. First, it was a retrospective study as the data were archival and was originally collected as part of standard admissions procedures at a CACREP accredited program. Therefore, this analysis was bound by the data that had been collected which could affect generalization of the study's findings to other programs with different admissions criteria. Similarly, there were limitations with one of the criterion used for admissions to this PCMP; namely undergraduate GPA. Poropat (2009) indicated that despite the widespread use of undergraduate GPA in admissions as well as research, there are concerns related to reliability and validity. Given that the undergraduate GPA scores used in this study were from participants who attended different academic institutions and majored in different academic disciplines, concerns such as grade inflation and ceiling effects were considered to be limitations. Second, participants were admitted to the program which likely contributed to a

restriction in the range of scores. Similarly, the data collected were from participants seeking admissions to a competitive MA program that typically receives 1.5 to 2 times more applicants than available positions. Therefore, it was possible that participants attempted to present themselves in a particularly favorable manner which may not have accurately reflected abilities and personality characteristics. Third, the participants all attended a PCMP at a medium sized university in the Rocky Mountain region. Accordingly, the findings may not generalize to other countries or regions of the United States.

Delimitations

The delimitations in the present study predominantly related to the design. It is estimated that there are more than 120 scales that can be scored and interpreted on the MMPI-2 (Weiner & Greene, 2008). The decision to delimit the MMPI-2 variables to the ten basic profile scores and three of the nine validity scales was based on the desire to keep the statistical model parsimonious and to ensure that the sample size was large enough to run a multiple regression analysis. Similarly, it was possible that other data submitted as part of the application process could have provided additional incremental validity such as demographic data, previous professional experience and letters of recommendation. The rationale for omitting these predictor variables was also based on the principle of parsimony as well as maintaining the ratio of predictor variables and participants such that a multiple regression analysis could be run. However, the main reason was that it was hypothesized that the three variables selected would explain the most variance. The desire to explain the maximum amount of variance was also the reason internship evaluations were selected rather than criteria such as years to graduate

and graduate verses not graduate. Lastly, comparing the amount of variance explained in the two models to determine which model was stronger was complicated by the fact that the dependent variables were not the same. Accordingly, results were reported with that caveat in mind.

Definition of Terms

Competency:

> "Professional competence is the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values and reflection in daily practice for the benefit of the individual and community being served" (Epstein & Hundert, 2002, p.226).

Criterion variable:

> "Variable being predicted or explained by the set of independent variables"(Hair, Anderson, Tatham & Black, 1995, p.81).

Incremental Validity:

> Incremental validity is defined as the degree to which a measure explains or predicts a phenomenon of interest, relative to other measures" (Haynes & Lench, 2003, p. 456)

Minnesota Multiphasic Personality Inventory-2 Validity and Scale

Validity Scales:

Lie (L) Scale:

> Designed to detect individuals who attempt to deliberately present themselves in a favorable light, but do so in a more obvious, unsophisticated manner (Graham, 2006).

Infrequency (F) Scale:

Designed to detect individuals who respond in a deviant or atypical manner (Graham, 2006).

Correction (K) Scale:

Designed to detect a more sophisticated attempt to deny or hide psychopathology or to present themselves in a more favorable light (Graham, 2006).

Clinical Scales:

Hypochondriasis, Scale 1 (Hs):

Developed to identify individuals who are preoccupied with their bodies and associated with fears of illness as well as disease. Although these fears are not considered to be delusional, they do tend to be persistent and individuals often will reject non-medical explanations (Graham, 2006).

Depression, Scale 2 (D):

Developed to identify individuals with symptomotology commonly associated with depression such as loss of hope, inhibition, seclusive and problems in thinking (Graham, 2006).

Hysteria, Scale 3 (Hy):

Developed to identify individuals who tend to react to stress in a hysterical manner. These reactions tend to arise when individuals feel overwhelmed and these individuals often lack insight regarding the underlying cause of symptoms (Graham, 2006).

Psychopathic Deviate, Scale 4 (Pd):

"Developed to identify patients diagnosed as psychopathic personality, asocial or amoral type" (Graham, 2006, p.73).

Masculinity-Femininity, Scale 5 (Mf):

Scale developed to measure how individuals identify with their respective gender role (Graham, 2006).

Paranoia, Scale 6 (Pa):

Developed to identify individuals who have symptoms associated with paranoia such as grandiose self concepts, rigid opinions and feelings of being persecuted (Graham, 2006).

Psychasthenia, Scale 7 (Pt):

Developed to identify individuals who have obsessive compulsive symptoms as demonstrated by "excessive doubts, compulsions, obsessions, and unreasonable fears" (Graham, 2006,p. 80).

Schizophrenia, Scale 8 (Sc):

Originally developed to identify individuals with schizophrenia but also identifies individuals who may be disorganized, disoriented and confused (Graham, 2006).

Hypomania, Scale 9 (Ma):

Identifies individuals with symptoms of hypomania as demonstrated by the following characteristics: "elevated mood, accelerated speech and motor activity, irritability, flight of ideas, and brief periods of depression" (Graham, 2006, p.86).

Social Introversion, Scale 0 (Si):

> "The scale was designed to assess a person's tendency to withdraw from social contacts and responsibilities" (Graham, 2006, p.88).

Predictive Validity

> "Predictive validity comprises statistical results indicating how accurately test scores can predict criterion scores at a later point in time" (Gall, Gall & Borg, 2007, p.198).

CHAPTER II

REVIEW OF THE LITERATURE

This chapter begins with a brief overview of the Minnesota Multiphasic Personality Inventory (MMPI) and the revised version, the MMPI-2. Empirical research using the MMPI/MMPI-2 to predict job and academic performance will then be presented. Focus will then shift to a general discussion on predictive and incremental validity and then to the context of how cognitive and non cognitive variables have been used to predict academic performance. Professional psychology then becomes central to the discussion where a review of current admissions criteria will be provided to highlight the level of satisfaction amongst directors of professional counseling training programs and to provide a context for understanding how important admission screening is in light of the rigorous graduate school training experience. Similarly, the importance of valid screening procedures is emphasized through a review of the literature to highlight the frequency of students who demonstrate difficulties with behaviors of competency, what kinds of problematic behaviors are most commonly identified and the ramifications for both the student and the professional counseling programs.

MMPI-2

The Minnesota Multiphasic Personality Inventory (MMPI) was created by Starke Hathaway, Ph.D., and J. Charnley McKinley, MD, and first introduced in 1943. They attempted to offer an instrument that could reliably and efficiently assist psychiatrists and

psychologists in assigning necessary psycho-diagnostic labels in clinical settings, as well as overcome limitations of existing personality assessments. Features such as empirically derived item selection and sound validity scales contributed to its instant success (Graham, 2006). After approximately forty years, the MMPI was re-standardized by Butcher et al. (1989, 2001) to update norms, better represent ethnicity and modernize some of the item content (Weiner & Greene, 2008). The inventory is now known as the MMPI-2 (MMPI-A for adolescents) and it has been received with the same esteem as its predecessor. Both the MMPI and the MMPI-2 have remained the most popular assessments for psychopathology research and clinical research as evidenced by the publication of more than 14,000 articles and books on the instruments and being the most widely used personality test throughout the world (Butcher, 2006; Butcher & Rouse, 1996; Graham, 2000).

The Minnesota Report: Revised Personnel Selection System, 3rd edition (Butcher, 2001) is a computerized MMPI-2 interpreted report marketed by Pearson Assessments. It generates profiles of scores on the standard validity and clinical scales, as well as numerous supplementary scales. Additional features include nongendered and gendered T scores and a narrative interpretation of scores. It offers six different reports specifically modified for certain public safety professions, and is said to provide in-depth, focused information to ease personnel selection. For instance, the Adjustment Rating Report provides the following ratings to assist in applicant selection: Openness to Evaluation, Social Facility, Addiction Potential, Stress Tolerance and Overall Adjustment. Although the breadth of information combined with the ease of interpretation could be enticing for a psychologist to utilize, it is important to underscore that the validity of the evaluations

generated by this report to predict job performance have not been published (Graham, 2006). Some researchers have published studies examining the concurrent validity of the computer assessment and other variables. For instance, Butcher (1988) compared the validity of scores between the report and experienced clinicians, and Muller and Bruno (1988) compared the MMPI-2 computer report to other variables, such as interview ratings and background checks. Both Butcher (1988) and Muller and Bruno (1988) found the MMPI-2 computer generated assessment to be commensurate with those other variables. However, there is still a need to empirically support the results of the MMPI-2 computer report to predict future job performance (Graham, 2006).

MMPI, MMPI-2 and Job Performance Research

Although the MMPI was originally intended for use with clinical populations, one of its more current applications has been to screen for sensitive occupations and for students in particular training programs. These occupations are associated with psychological factors that would significantly impact performance, including those that are high stress, involve personal risk and require personal responsibility (Butcher, 1985, 1991). In terms of appropriateness of use in this regard, Butcher, Ones and Cullen (2006) state that more than 570 books and articles have empirically supported the use of the MMPI/MMPI-2 in personnel and educational applications (Rouse & Butcher, 1995, as cited in Butcher, 2006, pp. 382-383). Literature on the ability of the MMPI/MMPI-2 to predict actual job performance, however, is more limited and complicated to determine.

There are several obstacles that make it difficult to determine the empirical validity of the MMPI-2 as a predictor of job and graduate school performance. First, if the MMPI-2 was used as part of the prescreening hiring process, it is likely that

applicants with profiles suggestive of serious pathology would not be hired. This significantly limits the ability to study deviant profiles and actual job performance, and restricts the range of scores (i.e., only scores in the normative range would be available). Similarly, if the MMPI-2 accurately identified applicants who would not be suitable for sensitive occupations and training programs, it would follow that there would be fewer individuals who would make it through the screening process and could possibly be identified later as exhibitors of poor job performance. Consequently, it can be challenging to obtain a large enough sample size of poor performers to make meaningful statistical interpretations (Hiatt & Hargrave, 1988). Therefore, it could be argued that the better the MMPI-2 performs at personnel screening the harder it could be to empirically validate the results.

A review of the literature demonstrates that the MMPI and the MMPI-2 have been used extensively in the selection of police officers (Butcher, 1979; Detrick, Chibnall & Rosso, 2001;; Weiss, Serafino, Serafino, Willson & Knoll, 1998; Weiss, Serafino, Serafino, Willson, Sarsany & Felton 1999). Numerous studies have been conducted on the validity of the MMPI-2 in predicting future job performance (Hargrave & Hiatt, 1987; Hargrave, Hiatt & Gaffney, 1988; Shusman, Inwald & Knatz, 1987). Hiatt and Hargrave (1988) conducted a multivariate analysis of variance to determine whether the MMPI could distinguish 53 "problem" (i.e., received serious disciplinary action) and 53 "nonproblem" officers. Results indicated that the MMPI profiles were significantly different for the two groups. The problem group had elevated T scores on Scales F, 5, 6, 9 and lower T scores on the L scale, which suggests that poor tolerance for frustration, impulsivity and hypersensitivity were personality characteristics that predicted poor

performance. This study also looked at T scores to determine the range at which profiles scores become predictive of job performance. Results indicated that 23 nonproblem officers and 39 problem officers had *T* scores at or above 70, which was the common cut-off score indicative of clinically significant pathology. Forty-eight of the scales in the profiles of the nonproblem officers had *T* scores at or above 65, and 95 scales in the profiles of the problem officers group had *T* scores at or above 65 (i.e., some officers had more than one scale score with elevations on their profile of scores). These findings suggest that it is possible that the prediction of an officer's job performance could be improved by utilizing adjusted cut-off *T* scores.

It is notable that all of the officers in Hiatt and Hargrave's (1988) study were prescreened on several measures, including the MMPI, and hired, despite having profiles that indicated psychological concerns. This is significant because they overcame a common obstacle in predictive validity research—namely, range restriction. Their findings demonstrated that even with range restriction, personality characteristics were still able to differentiate between the two groups and provided useful information about job performance in a population of officers who have scores in the clinically significant range. Another strength of this study was that Hiatt and Hargrave (1988) combined departments to have a big enough sample of problem officers, and matched these officers with nonproblem officers on all other covariates. This allowed for more meaningful interpretations because they accounted for moderator effects, such as age and gender.

In a similar investigation of police officer performance, Shusman et al. (1987) indicated that the validity Scale K and Scale 4 (Pd) best discriminated "good" from "poor" job performance as indicated by absences, lateness and disciplinary interviews.

Weiss et al. (1998) researched the ability of the MMPI-2 to predict job performance ratings and the continuation of employment amongst 32 recently hired police officers. A strong correlation was found between high raw scores on the Paranoia Obvious (Pa) scale and Paranoia Subtle scale and job removal ($p<.05$). Additionally, high scores on the Paranoia Obvious (Pa) scale were found to predict poor performance ratings. Given that these officers were administered the MMPI-2 as part of the hiring process, the authors discussed the likelihood that it took time for subtle paranoid tendencies to arise, and concluded that the Pa scales have predictive validity for supervisory ratings and retention amongst police officer applicants. Lastly, Chibnall and Detrick (2003) investigated the incremental validity of the MMPI-2, the Revised NEO Personality Inventory and the Inwald Personality Inventory with police academy performance as the criterion. Results indicated that all inventories significantly contributed to the prediction of academic performance.

The literature also indicates that some airline carriers conduct psychological evaluations as part of their application process (Butcher, 2002). However, there is still a need to empirically validate whether it can predict future performance as a pilot. Butcher (1994) studied 437 airline pilots and reported that the MMPI-2 norms were better suited for nonclinical populations than the original MMPI. The study also found that applicants are more likely to respond defensively. Therefore, interpretation of high elevations on Scale L and Scale K should be interpreted with that caveat in mind. In fact, Butcher (2002) explained that these scales were not developed for the purpose of personnel screening, therefore underscoring the limited utility of using these validity indicators in that capacity.

In a study of nurses who were promoted to supervisor, Kelly (1974) found that personality factors measured by the MMPI, CPI, 16 PF and EPPS differentiated those who were promoted from those who were not. A review of other studies conducted to confirm the efficacy of the MMPI and MMPI-2 for personnel screening of sensitive occupations include the following: nuclear power plant employees (Dunnette, Bownas & Bosshardt, 1981) military personnel (Callan, 1997; Carbone, Cigrang, Todd and Fielder, 1999), clergy (Banks, Mooney, Mucowski & Williams, 1984; Celmer & Winer, 1990; Jansen, Bonk & Garvey, 1973), and psychiatric residents (Garetz & Anderson, 1973).

MMPI, MMPI-2 and Academic Performance Research

Butcher and Rouse (1996) maintained that the MMPI-2 does detect individuals who are likely to experience emotional maladjustment in a stressful academic environment. Although recent research to support that assertion is sparse, there is evidence that personality characteristics as measured by the MMPI and MMPI-2 can predict academic performance. Lachar (1974) studied a freshman class at the United States Air Force Academy (N=1,389). Cadets were administered the MMPI when they entered the academy and then followed up on, including those that dropped-out after the initial period of adjustment. Findings indicated that the rates of attrition and emotional maladjustment were significantly higher for the group of cadets identified as being "high risk" when screened with the MMPI.

Strupp and Bloxom (1975) compared male undergraduate college students with T scores above 60 on Scale 2 (Depression) and Scale 7 (Psychasthenia) with a random group of undergraduate students over the course of eight-and-a-half years. Results indicated that the students with high scores on Scale 2 and Scale 7 had more emotional

and behavioral problems, as well as longstanding, negative concerns related to academic, personal and occupational performance. In other words, they struggled more than the comparison group of college students and in the subsequent establishment of a family and occupation. A similar study by Davis and Widseth (1978) supported the findings by Strupp and Bloxom (1975). These researchers found that amongst 755 entering male college students, 21% had T scores above 60 on Scale 2 and Scale 7. These students were significantly more likely to take a leave of absence, have personal problems (i.e., "internal psychological problems") and seek mental health services.

Kodman (1984) analyzed the mean group differences of 100 Summa Cum Laude university students (50 male and 50 female) with 100 other senior university students, who were matched by race, gender, age and education. Fisher t-test results indicated that there were statistically significant differences for the two groups of males on the following scales: Hysteria (Hs), Psychopathic Deviate (Pd), Masculinity/Femininity (Mf), Psychasthenia (Pt), Schizophrenia (Sc) and Social Inversion (Si). The same scales were found to be statistically different for females, with the exception of the Si Scale. As a whole, the honors group scored higher than the control group on all the subscales found to be significant. Additionally, the mean subscale scores for the honors group peaked on subscales 7 (Pt) and 8 (Sc). Although Kodman stated that the Summa Cum Laude sample "scored more in the pathological direction," only the mean subscale T scores were reported (p.2). Nonetheless, the findings do contribute to a better understanding of an MMPI profile and frequency scores for high achieving undergraduate students.

When the MMPI was revised, Butcher, Graham, Dahlstrom and Bowman (1990) empirically validated that the MMPI-2 was an appropriate instrument to administer to

college students by comparing 3 validity scale and 10 clinical scale scores of the normative sample to a sample of 1,312 college students from four different and regionally distinct universities in the United States. Based on the reliability of scale scores, frequency distributions and test-retest correlations, Butcher et al. concluded, "Researchers and clinicians using the MMPI-2 with college students can be assured that the MMPI-2 norms are appropriate for this population" (p. 14).

Using the MMPI-2, King and Bailly (2002) analyzed the relationship between personality and undergraduate GPA. Pearson product moment correlations indicated that elevations (T>59) on the following scales were associated with "poor academic performance" (i.e., GPA): F, Pd, Sc, and Ma. The number of elevated scales did not contribute to the prediction of academic performance, and a stepwise multiple regression analysis indicated that the combination of elevated scales (e.g., code type) did not explain more variance than single scale scores.

Compton (2009) conducted a similar study trying to predict first-year cumulative GPA of undergraduate students (N= 571) using 8 validity scales and 10 clinical scales on the MMPI-2. After calculating Pearson correlation coefficients for MMPI-2 scale scores and GPA, only the following scales were regressed on the criterion of GPA: Infrequency (F), Infrequency Back (Fb), Psychopathic Deviate (Pd), Paranoia (Pa), Hypomania (Ma) and Social Introversion (Si). Results indicated that the final model accounted for 10.6% of the variance in GPA scores. Notably, the Ma scale accounted for 7.5% in that final model. Given that Compton concluded that the MMPI-2 is not a good predictor of GPA, it is worthwhile to briefly examine the limitations in that study and the subsequent conclusions. First, the sample was comprised of students enrolled in a private, vocational,

religious undergraduate institution where the majority of students are professionals who were changing careers (i.e., mean age of sample=33 years old). Not only do these factors negatively impact the ability to generalize the results to a typical undergraduate college population, but it is also likely that there were significant moderator effects left unexamined. Second, despite the number of scales on the MMPI-2, conceptually, the regression model only used one variable, personality, to predict a multidimensional criterion, academic performance. It is possible that if Compton had added demographic or cognitive independent variables to explain additional unique variance, an increment of 10.7% of variance accounted for by personality characteristics might have been considered significant to the overall model.

Predictive Validity, Incremental Validity and Criterion Issues

There is a lack of consensus in the literature as to what are the most valid predictive measures to utilize for admissions criteria to graduate psychology programs, as well as debate regarding a definition for the criterion of "success." More recently, research on academic performance, in general, has focused on incremental validity rather than approaching prediction from the standpoint of one independent variable being superior to another. "Incremental validity is defined as the degree to which a measure explains or predicts a phenomenon of interest, relative to other measures" (Haynes & Lench, 2003, p. 456). In applied psychology, improved prediction can mean several things, such as increased specificity or the power and ability to make decisions. The emphasis is on the extent to which a variable adds to the prediction model beyond what could have been predicted with other variables (Hunsley, 2003), or simply utilizing base rate data (Hunsley & Meyer, 2003).

In addition to distinguishing that new data can strengthen an existing statistical model, incremental validity research can also determine how well a variable accounts for outcome variance in comparison to the other measures (Haynes & Lench, 2003). It can provide evidence of whether multiple sources of data should be jointly considered, irrespective of whether individually the variable is known to contribute to the prediction of a criterion. Hunsley (2003) stated, "Any scientific justification for obtaining and combining multiple sources of data can only come from research evidence in which unique and shared variance among measures is directly examined with respect to a criterion of interest" (p. 443). Therefore, incremental validity research can assist in the development of a more parsimonious model because it can determine if the addition of multiple predictor variables improves the statistical model beyond what variance could have been explained by fewer variables. Utility, finally, is also an important consideration that can be addressed through incremental validity research because it can be used as evidence to support or refute the use of additional variables based on elements such as time and cost (Hunsley, 2003).

Hunsley and Meyer (2003) stated, "The focus on this type of research is on the value of adding new test data into a statistical equation, generally on the basis of regression analysis, in order to predict a criterion" (p. 448). For illustrative purposes, some incremental validity research using the MMPI-2 is provided. Wetzler, Khadivi and Moser (1998) performed a stepwise hierarchical regression analysis, and found that the content scales on the MMPI-2 provided incremental validity over the 13 clinical and validity scales for the diagnosis of depressive disorders and psychotic disorders in a sample of 264 psychiatric inpatients. Using hierarchical regression analyses, Wygant,

Sellbom, Graham and Schenk (2006) found that the Personality Psychopathology-Five (PSY-5) scales had incremental validity over the clinical and content scales on the MMPI-2 in the prediction of personality disorders. Dao, Prevatt and Horne (2008) tested the incremental validity of the Rorschach Perceptual Thinking Index (PTI; Exner, 2003), clinical Scale 6 (Pa), Scale 8 (Sc), Bizarre Mentation (BIZ) and the Goldberg Index on the MMPI-2 in the differentiation of nonpsychotic and psychotic inpatients. Results indicated that the PTI was able to add incremental validity when the MMPI-2 variables were entered into the hierarchical regression analysis first. However, when the PTI was entered into the analysis first, the MMPI-2 scales did not add to the predictive strength of the model. Lastly, Miller (2009) conducted hierarchical multiple regression analyses to test whether personality characteristics, as measured by the PSY-5 Scales on the MMPI-2, could add incremental validity beyond demographic, contextual and historical variables in the prediction of risk-taking behaviors (i.e., smoking, drinking alcohol and sexual behavior) in college. Results indicated that the PSY-5 Scales accounted for the most variance on all outcomes, except the influence of peers regressed on risky consumption of alcohol. Prior to a review of incremental validity research on the prediction of individual differences in academic outcomes, it is important to briefly address important considerations in this line of inquiry. Outcomes are impacted by the regression model (e.g., stepwise verses hierarchical multiple regression) used by the researcher and interpretation of the data should be made with that caveat in mind. For example, in hierarchical multiple regression models the order in which the data is entered in the regression analysis will determine which variable is assigned the shared variance. Another factor that can influence the amount of incremental validity is the reliability of

the predictive measure, because the more reliable one measure is over another, the more incremental validity contributed to the results. Lastly, utility is an important consideration because the amount of incremental validity a predictor variable needs to account for to be considered significant is researcher driven (i.e., there is not a statistical set point) (Hair, Anderson, Tatham & Black, 1995; Hunsley & Meyer, 2003). For example, if a program is experiencing budget cuts and needs to justify the use of a measure, it may require a higher percentage of variance accounted for than research that is not concerned about cost.

Predictive Validity: Intelligence and Academic Success

There is strong support in the literature for the inclusion of cognitive variables when building a statistical model to predict academic performance. In fact, there is a general consensus that intelligence is a robust predictor of academic success (Busato et al., 2000; Di Fabio & Palazzeschi, 2009; Farsides & Woodfield, 2003; Furnham & Monsen, 2009; Trapmann et al., 2007). The Task Force on Intelligence, established by the American Psychological Association in 1995, concluded that intelligence test scores predict school achievement reasonably well with correlations of .50 for grade point average and .55 for total number of years of education (Neisser, Boodoo, Bouchard, Boykin, Brody, Ceci, Halpern, Loehlin, Perloff, Sternberg, & Urbina, 1996, p. 96). In terms of the predictive validity of school grades and college GPA, a meta-analysis conducted by Burton and Ramist (2001) reported a correlation of .42. Likewise, intelligence scores correlate with success outside of the academic domain in areas such as occupational status, and show negative correlations with behaviors such as criminal activity. It is notable that the high correlation between intelligence and occupational

status remains significant even when effects of education and family background are statistically controlled (Neisser et al., 1996). For example, to test the incremental validity of five psychometric instruments, Furnham, Monsen and Ahmetoglu (2009) found that intelligence tests can reliably predict exam scores for high school students. In fact, after controlling for intelligence, approach to learning and personality accounted for very little of the variance. In light of these findings, it is important to underscore that the age of the participants was likely significant, in that the range of scores were more variable than what would be found amongst students already admitted to university and, therefore, accounted for more variance. Nonetheless, the literature suggests that even when range restriction and cognitive measures are used to predict college GPA, approximately 25% of the variance can be explained (Trapmann et al., 2007).

Incremental Validity: Personality and Academic Performance

Personality measures provide incremental validity in the prediction of academic and work performance because of low to zero correlations with cognitive ability (Kuncel & Hezlett, 2010; Zeidner & Matthews, 2000). Personality is also an important variable to consider in performance research and predictive statistical models because cognitive tests of ability and aptitude measure what an individual is capable of doing. However, personality characteristics can provide information about how an individual is likely to perform (Furnham, Monsen & Ahmetoglu, 2009).

Over the past twenty years, empirical research on the relationship between personality and academic performance has burgeoned, and a review of the literature has demonstrated that researchers have predominantly used the Big Five Factor Model of personality as the taxonomy for classifying personality dimensions.

> The most coherent framework and consistent results have derived from studies on the five-factor model (FFM) or *Big Five* personality traits, which asserts individual differences in normal behavior should be classified in terms of five orthogonal or independent dimensions, namely Neuroticism, Extraversion, Openness to Experience, Agreeableness and Conscientiousness. These dimensions reflect individual differences in stable dispositions and preferences that determine each individual's characteristic patterns of thought, emotionality and behavior; they represent aggregated measures of individuals' behavior and can be assessed through self or other reports (Furnham, Monsen &Ahmetoglu, 2009, p. 770).

For example, personality traits as measured with instruments based on the Five Factor Model have been found to add incremental validity in predicting a range of academic outcomes, such as final grades and grade-point average at various levels of education (Chamorro-Premuzic & Furnham, 2003; Lievens, Dilchert &Ones, 2009; Poropot, 2009).

Ozer and Benet-Martinez (2006) indicated that the trend of using the FFM in predictive and incremental validity research has advanced the understanding of how personality can account for individual differences in academic performance. For example, numerous studies have shown that personality characteristics account for more variance in performance as individuals progress academically to higher levels of education. Additionally, utilizing the FFM as the classification system has conformed the research by creating a universal language to communicate information about personality constructs (Furnham, Monsen & Ahmetoglu, 2009). That being said, it would be misleading to assume that the volume of research using the FFM, and not other valid, reliable and objective personality instruments, reflects a consensus in the field of professional psychology that the FFM is the best method for understanding and measuring dimensions of personality. Chamorro-Premuzic and Furnham (2006) addressed this trend and explained, "The snowball effect regarding the consensus on the Big Five is indicative of a self-fulfilling prophecy: The more researchers agree on the

validity of the Big Five, the more research is conducted to validate the Big Five" (p. 256). Therefore, it is important to underscore that the reason research using the FFM will dominate this section of the literature review is because there is a significant lack of recent research that has used other personality assessments, such as the MMPI-2.

In a meta-analysis of 58 studies on prediction of academic performance from measures of the Big Five personality factors from 1980 to 2006, Trapmann et al. (2007) examined three different criteria: GPA, retention and academic satisfaction. They concluded that the predictive strength of personality depends on the criterion with which it is measured. Only one of the five factors, Conscientiousness, was found to predict college grade-point average (7.2% of the variance). However, a hierarchical regression analysis revealed that when conscientiousness was added to the model after high school grades, only an additional increment of 1.8% of the variance was explained. Meta-analytic results also indicated that the five factors did not predict retention. In terms of academic satisfaction, extraversion correlated with satisfaction and neuroticism with negative results.

Different results were found by Poropat (2009), who conducted the most comprehensive investigation of personality-academic performance relationships to date, and concluded, "Personality should take a more prominent place in future theories of academic performance and not merely as an adjunct to intelligence" (p. 333). This assertion was manifest from a meta-analysis that consisted of a cumulative sample size of more than 70,000, and included analyses of primary, secondary and tertiary levels of education, with the emphasis on the latter. Personality measures were based on the FFM and the criterion was GPA. Findings indicated that three of those characteristics,

Agreeableness, Conscientiousness and Openness, significantly correlated with academic performance. In fact, Conscientiousness (21% of variance explained) and intelligence (22% of variance explained) were found to have commensurate levels of validity in both secondary and tertiary levels of education.

Noftle and Robins (2007) also concluded that personality traits provided incremental validity after analyzing the relationship between the Big Five personality traits and academic outcomes. Conscientiousness was the most robust predictor of high school and college GPA and Openness was found to be the strongest predictor of verbal SAT scores. It is notable that these findings not only replicated previous research (e.g., DeRaad & Schouwenburg, 1996; Farsides & Woodfield, 2003; Loundsbury, Sundstrom, Loveland & Gibson, 2003) but also replicated within this study by four independent samples that used four different personality inventories.

Lievens, Dilchert and Ones (2009) longitudinal study investigating the predictive-criterion relationship between personality and academic performance is the last to be reviewed in this category, and has been selected because the findings demonstrate how important it is to have well-defined criterion in the model, as well as how significant a factor-restricted range and other moderator effects are in tertiary education. The entire 1997 medical school cohort in Belgium were the focus of a study aimed at identifying whether validity changes over time, and found that the predictive validity of personality traits increased as students progressed through their seven-year medical programs. This is especially significant for admissions committees in programs where application of knowledge does not take place until the later in the process, such as counseling practicum and internship, where knowledge turns into applied practice.

Predictive Validity: Performance in Professional Psychology Graduate Programs

It has been established that cognitive and personality variables do provide increments of validity in the prediction of individual differences in academic outcomes. Likewise, it has been documented that, as individuals progress to higher levels of education, increments of validity from cognitive factors decrease and increments of validity from personality increase due to range restriction and the emphasis on knowledge translating into applied practice (Lievens, Dilchert & Ones, 2009; Trapmann et al., 2007). The focus of this study will now shift to a review of the literature pertaining specifically to academic success in the field of psychology to delineate what variables have been validated to account for variance in individual academic outcomes.

Empirical support for the utility of using GRE as a variable in the prediction model for success in graduate psychology programs is mixed. Several researchers have tested GRE using graduate verses non-graduate as the criterion of success. For instance, a five-year study conducted by Hirschberg and Itkin (1978) at the University of Illinois at Urbana-Champaign, showed the success of all graduate students admitted to the Department of Psychology from 1965 through 1970. The researchers found that none of the standard admissions variables, such as GPA and GRE, were predictive of who would graduate with a PhD. Holmes and Beishline (1996) found that GRE identified false negatives in terms of prediction of students who completed their doctorate in psychology. In other words, there were students whose scores on GRE predicted they would not be able to complete their doctorate, but who went on to graduate. However, Holmes (2003) responded to the outcome of that research by highlighting that the students who scored

lower than 1000 on GRE did produce false negatives in terms of completion of doctoral programs. Scores higher than 1100 did, however, accurately predict completion of graduate study in psychology. Accordingly, Holmes concluded that the reason virtually all doctorate programs in psychology continue to use GRE is because there are simply more applicants than available positions, and students with higher GRE scores have a higher chance of completion.

Widening the scope for the criterion of success, Sternberg and Williams (1997) investigated the empirical validity of GRE to predict the following performance criteria in a graduate psychology program at Yale University: grade point average after the first and second year in the program and professor ratings of the student's abilities in tasks such as analytical, creative, practical, research and teaching. Results indicated a correlation of .17 for the combined median scores on GRE (i.e., Verbal, Quantitative, Analytical and GRE Advanced test in Psychology) for men and women using first- and second-year grades. When analyzed separately, GRE did not produce significant correlations with second-year GPA. The GRE Analytic scale was the only variable able to predict professor ratings of student's analytic abilities, and this correlation (.26) was only significant for the men in the sample.

Sternberg and Williams explained that their empirical study was inspired by the "triarchic theory of human intelligence" (Sternberg, 1985):

> The triarchic theory distinguishes among academic-analytical, synthetic-creative, and practical-contextual aspects of human abilities…. The theory also suggests that practical and especially creative abilities will be critical for performance as a psychological researcher or practitioner and that although analytical abilities will also be important, they will not hold any privileged position (Sternberg & Williams, 1997, pp. 633-634).

Given that GRE was not able to predict what Sternberg and Williams consider important criteria of success, they suggest that alternative predictor variables should be researched to assist in the identification of students best suited for the field. This is significant because it underscores the necessity of not only tailoring criteria of success to the specific objectives of the training program, but also of considering multiple sources of data when building a statistical model to predict criteria of success in professional psychology programs.

A five-year meta-analysis by Goldberg and Alliger (1992) that focused on the predictor-criterion relationship between cognitive variables and academic outcomes in graduate psychology/counseling programs suggested that GRE does not significantly contribute to the prediction of success. In order to be included in the meta-analysis, the previous research had to include at least one of the three measures on GRE (i.e., Quantitative, Verbal or Advanced Psychology) as a predictor variable. The correlations between the GRE Quantitative portion and the comprehensive exams was r=.37. GRE was not found to be a valid predictor of student success as measured by graduate GPA because only 2% of the variance in graduate GPA was accounted for in the analysis (Goldberg & Alliger, 1992). However, Chernyshenko and Ones (1999) underscored that Goldberg and Alliger did not correct for range restriction, and they re-analyzed the same data, which produced significantly different results. “Correction for range restriction allows one to estimate the operational validity of the GRE scores in selection situations” (Chernyshenko & Ones, 1999, p. 953). The highest mean correlations, after correcting for range restriction, were found between the GRE Verbal score and the comprehensive exam scores (r=.70). The smallest criterion-related validity score was between GRE

quantitative scores and graduate GPA (r=.35). “The magnitude of the reported GRE validity coefficients showed that high-scoring individuals have greater probability of succeeding in psychology graduate programs (or at least in passing comprehensive exams and receiving high graduate GPA) than low scoring individuals” (p. 959). Therefore, Chernyshenko and Ones provided solid rationale for including cognitive measures of achievement in academic prediction models once effects for range restriction had been accounted.

Personality characteristics have also been empirically validated as predictors of success in the field of psychology. Chamorro-PreMuzic & Furnham (2003) conducted two longitudinal studies on psychology undergraduates to test the predictability of academic performance by various independent variables. It was found that both the NEO Five- Factor Inventory- Revised (NEO- FFI) and Eyesenck Personality Questionnaire-Revised (EPQ-R) could predict the academic performance of undergraduate students (i.e., psychology majors) after three years. In the first sample, the NEO-FFI was the personality measure and was found to be a better predictor than academic variables (i.e., 10% of the unique variance on exam scores), especially the Neuroticism and Conscientious factors. In the second sample, 17% of the unique variance was accounted for by the EPQ-R.

Given the lack of consensus on the definition of success in professional psychology programs, the literature was reviewed to identify criterions other than graduate GPA. Smaby, Maddux, Richmond, Lepkowski and Packerman (2005) conducted a study of 80 master’s level counseling students to specifically assess whether academic variables at the time of admissions (GPA, GRE Verbal and GRE Quantitative

scores) could effectively predict the following: knowledge as indicated by the Counselor Preparation Comprehensive Examination (CPCE); personal development as indicated by the Counselor Skills and Personal Development Rating Form (CSPD-RF); and mastery of counseling skills as indicated by the Skilled Counseling Scale. Multiple regression analysis supported the hypothesis that GRE Verbal and undergraduate GPA could significantly predict overall performance on comprehensive exams. A significant model was not found to predict personal development; however, GRE Verbal scores and GPA were predictive of counseling skills acquisition. Given these findings, the authors recommended that other methods for assessing personal development at admissions be researched.

Using multiple regression analysis, Costanzo and Philpott (1986) searched for reliable predictors of psychotherapeutic talent. Their model included the following independent variables: Therapy Process Analysis Task and the Social Interpretation Task ; the Adjective Checklist and Allport's Study of Values to measure personality, demographics of age, sex and previous counseling experience; and examination scores from an "upper division" undergraduate course in clinical psychology to measure academic achievement. The Therapy Process Analysis Task and the Social Interpretation Task were utilized to measure interpersonal intelligence. The former consisted of participants watching a video-taped counseling session. They were then asked to write an essay identifying key points in the session and make recommendations, which was blind scored on a Likert-type scale by a professor and a graduate assistant. The latter task had participants view a 30 minute video-tape of 20 different interpersonal interactions which participants were asked to interpret. The Adjective Checklist and Allport's Study of

Values was administered to measure personality traits. The former is a self-administered checklist of 300 adjectives that describe behavior. Costanzo and Philpott only used the subscales they assessed to be conceptually similar to the construct of therapeutic talent, namely Communality, Intraceptiveness, Nurturance, Masculine Orientation and Feminine Orientation. "Allport's Study of Values assesses the relative importance of six common value orientations: religious, economic, social, theoretical, political, and aesthetic, for an individual" (Costanzo & Philpott, 1986, p.365). The criterion of therapeutic talent was measured by scores on the Group Assessment of Interpersonal Traits (GAIT) that resulted in a global measure of therapeutic talent. This score was calculated from the ratings of two licensed clinical psychologists and a masters level counselor based on participants performance in a simulated initial counseling session.

> The GAIT sequences were presented in random order and rated along the following dimensions: therapeutic talent, empathy, acceptance, nonverbal attentiveness, insight, emotional-affective orientation, intellectual-cognitive orientation, and rapport between understander and discloser (Costanzo & Philpott, 1986, p.365).

Forty-two percent of the variance in therapeutic talent was explained by the model, meaning that 42% of the variability in therapeutic talent was predicted by knowing the scores of the independent variables. The strongest predictors were the two measures of interpersonal intelligence. Based on these results, the authors underscore the importance of assessing interpersonal intelligence as part of admissions. While this suggestion may be prudent, one of the weaknesses of this study was the use of examination scores from an "upper-level" undergraduate course in clinical psychology. Given that the authors did not provide any information about the course or even provide a title, the study cannot be replicated. Moreover, the amount of incremental validity a

variable can provide is affected by the reliability of that predictive measure, and it is questionable whether the scores from an exam in one college course could be considered valid and reliable estimates of academic achievement by which to measure other variables against.

Daehnert and Carter (1987) conducted a more comprehensive study that looked for precise measures to predict those most suited for the profession of psychology. The authors maintained that traditional admission criteria did not adequately address the multiple training objectives of a professional, psychology training program, and they suggested that admissions criteria should be driven by the programs specific training objectives. Their sample consisted of 192 students who were admitted to either a master's program or an APA-accredited doctoral level program over seven consecutive years. The following predictor variables were utilized: academic and achievement measures (GRE and undergraduate GPA), personality variables (MMPI and Strong Vocational Interest Blank), indices based on letters of recommendation, and historical biological/educational variables (quality and type of undergraduate institution, age at admissions, gender, marital status, and presence of an advanced degree). According to Daehnert and Carter the most significant finding was that personality variables, as measured by the MMPI, were useful predictors for internship evaluations. Specifically, Pearson product-moment correlations had a strong relationship between the internship subcategory "Knowledge of Personality Theory" and Scale 7, Psychasthenia (.56), as well as the K Scale Defensiveness (-.51). "Knowledge of Personality and Counseling Theory" was negatively correlated with Scale 4 Psychopathic Deviate (-.21) as was Scale 8 Schizophrenia (-.41) and the K Scale (-.33) with "Knowledge of Psychopathology". Lastly, high scores on

internship ratings of Responsibility (.35) and Therapeutic Relationship skills (.34) were positively correlated with high scores on Scale 6 Paranoia.

Given that the correlation between personality factors and internship evaluations were the most significant findings of the study, and that personality factors did not predict academic performance, Daehnert and Carter (1987) discussed two important implications. First, they asserted that internship evaluations are "ideal criteria" for measuring graduate school success, and, since the MMPI was the only instrument that could predict this criterion, it follows that Pre-Admissions criteria should include measures of interpersonal and intrapersonal functioning. Second, they stated that because the MMPI did not predict academic performance, it is an essential predictor of graduate success in psychology, but insufficient by itself. The latter assertion may be true; however, it also highlights one of the shortcomings of this study. Although the number of predictor and criterion variables used in this study was impressive, the small sample size precluded the use of multiple regression analysis. Unfortunately, the Pearson product-moment correlations and factor analysis did not provide an adequate or meaningful understanding of the relationship among the numerous variables. Therefore, replication of this study with a large enough sample to run a multiple regression analysis would be valuable to better understand how much unique and shared variance the predictors could account for in graduate psychology student performance.

Professional Counseling Graduate Programs: Current Admissions Procedures

Professional psychology admissions committees contend with the challenging task of choosing applicants from an often well-qualified, homogeneous pool in terms of

academic history and prior professional experience. Although admissions criteria for graduate study vary amongst programs, there are some definite commonalities. Norcross, Kohout and Wicherski (2005) conducted an analysis of 1,970 distinct graduate programs and 601 departments in graduate psychology, and reported descriptive data about graduate study in psychology. In 2003-2004, master's level–only programs rated the following as being the most important criteria for admissions decisions: GPA, letters of recommendation, and personal statements. Doctoral programs rated letters of recommendation, personal statements, GPA, interview, research experience and GRE scores as being the most important criteria. Fifty-seven percent of master's programs and 74% of doctoral programs in their study required GRE. Eighty-six percent of master's programs and 82% of doctoral programs required the prospective students GPA to make admissions decisions.

In terms of methodology, Norcross, Kohout and Wicherski (2005) asked participants to rate nine admissions criteria on a four-point scale. It is notable that an objective personality assessment, such as the MMPI-2 or NEO-FFI, was not one of the nine criteria offered. Rather, the research suggested that if personality or character was considered by programs, it was done so subjectively through criteria such as personal interviews or letters of reference. Similarly, Johnson and Campbell (2004) reported that directors of admissions for doctoral psychology programs used letters of recommendation, admissions interviews and personal statements during the screening process to assess the character and fitness (emotional/mental health) of applicants. Accordingly, review of the literature was conducted to discover whether the personality characteristics of applicants were considered to be important admissions criteria in

graduate psychology programs, and to assess the level of satisfaction with current admissions screening procedures. Johnson and Cambell (2004) focused on the character and fitness of doctoral psychology students through the lenses of directors of clinical training (DCT). On a 5-point Likert-type scale (1=*not at all concerned*, 5=*extremely concerned*), DCT endorsed being significantly concerned about both moral character (mean=4.02, SD =.95) and psychological fitness (mean=4.36, SD=.74). According to Johnson and Campbell (2004), "Character represents the honesty and integrity with which a person deals with others…. Psychological fitness has to do with emotional/mental stability" (p. 406). The authors underscore that one of the problems faced by DCTs is that there is not a consensus as to how to detect character and fitness.

> The most salient difficulty with efforts to evaluate applicant and student character and fitness in psychology doctoral training programs is the absence of any valid tool or protocol designed for this purpose. Currently, no research in psychology demonstrates the efficacy of a screening approach to character and fitness (Johnson & Campbell, 2004, p. 406).

Pre-Admissions psychological testing was endorsed by a few DCTs, and the authors suggested objective measures as a viable screening procedure. However, the authors noted that there is no empirical data to support the predictive validity of using these measures as part of admissions.

Leverett-Main (2004) conducted a similar investigation from the perspective of program directors in CACREP-accredited counselor education graduate programs. She stated,

> Graduate programs in counselor education need a reliable methodology that can be used to effectively screen applicants for admission. Such screening procedures are intended to ensure capable individuals who evidence potential for success both academically and in the field are identified for admission to counselor education programs…. It is, therefore, imperative that counselor education admissions committees find a variety of measures that effectively assess a

> student's ability to deal with the multiple demands of an accredited graduate program (Leverett-Main, 2004, pp. 207-208).

Sixty percent of the 361 CACREP-accredited counselor education programs participated in the study. Personal interviews were endorsed by program directors as being the most efficacious means to screen applicants, and 86% endorsed practicum/internship evaluations as the best measure of success in the program. Overall, program directors indicated that they were satisfied with current admissions procedures as evidenced by satisfaction rating, which ranged from 72% to 92%.

The Training Experience

Given that training in counseling psychology is a rigorous, challenging and often stressful experience, it is imperative that admissions committees screen for applicants best suited to contend with the multiple demands. Accordingly, several scholars have reviewed literature that is relevant to the training experience of and mental health of graduate psychology students to better understand common concerns, and to see if there is evidence of additional Pre-Admissions screening methods. Kuyken, Peters, Power and Lavender (1998) measured adaptation to training amongst 183 clinical psychology trainees and found that the mean level of adaptation was in the average range; however, significant levels of anxiety, depression, work adjustment problems and self-esteem problems were indicated amongst more than 25% of the trainees.

Brooks, Holtum and Lavender (2002) furthered that research by exploring how course-related experiences and individual personality factors impacted clinical psychology students' ability to adapt. The overall adjustment score was measured with the Millon Index of Personality Styles (MIPS) (Millon, 1994) and results indicated that the trainees' (N=364) mean overall personality adjustment score was significantly better

than the normative sample. However, the trainees recorded experiencing more symptoms on the depression, anxiety and self-esteem scales, but these scores were still well above the level indicative of poor adaptation. Likewise, frequency data demonstrated that 23% of the sample reported significant self-esteem problems, 18% had significant anxiety problems, 14% endorsed significant depression and 8% had significant work adjustment problems. Overall, 41% endorsed at least one of the four problems previously identified, and approximately one-third of the total sample endorsed significant substance abuse problems.

Another study by Kuyken, Peters, Power and Lavender (2003) found that among 183 clinical psychology trainees, difficulties such as coping with and adaptation to graduate school, not only increased in terms of the number of distressed students over the course of two years, but also the severity of the problems increased. The authors provided the following reasons why it would be prudent for trainees with adaptation problems to be identified:

> (1) trainees who appear to be functioning well overall may be experiencing significant problems in one or more areas, (2) a significant proportion of trainees experience enduring problems, (3) often problems tend to be of the internalizing type, and therefore not readily apparent, (4) some of the identified problems may have considerable personal and professional ramifications and (5) difficulties tend to increase during training, both in terms of numbers of trainees and severity of self-reported problems (Kuyken et al., 2003, p. 52).

Accordingly, Kuyken et al. (2003) asserted that an "obvious point of prevention" (p. 52) would be in the selection process of candidates.

Smith, Robinson and Young (2007) approached this topic from the perspective of wellness and focused on master's level counseling students entering the programs. Two-hundred-four students from nine different CACREP accredited programs were given the

Five Factor Wellness Evaluation of Lifestyle (5F-Wel) test. Overall, the students scored better (i.e., indicating higher level of wellness) than the normative population; however, the Total Score for 10.7% of the student sample endorsed symptoms of psychological distress commensurate with levels found in clinical populations. Accordingly, Smith et al. concluded, "Given that 10.7% of the sample scored above the clinical cutoff score for Psychological Distress, future research needs to examine the skills and success levels of students representing this population" (p. 105).

Competency

Wilkerson (2006) maintains that there is an inadequate supply of empirical research on the subject of trainee impairment and highlights how problematic this is by providing data to support the global assertion that nearly all professional training programs are confronted with this issue. Kerl, Garcia, McCullough and Maxwell (2002) examined issues related to screening applicants for counseling psychology programs and asserted the following:

> Although counseling programs use admissions criteria to select students who, they believe, will be successful in their programs, it is unrealistic to rely entirely on screening procedures during the admissions process to identify students who do not have the necessary personal characteristics to become competent counselors. Personal qualities of competent counselors include the capacity for empathy, genuineness, acceptance, access to and appropriate sharing of feelings, giving and receiving feedback effectively, honesty, and establishing and maintaining relationships. Individuals who apply to counseling programs may have issues or characteristics that are not readily identifiable through admissions procedures, but these can be issues, such as biases, lack of impulse control or information-processing deficits, that may impair the ability to practice effectively (p. 322).

The literature was reviewed to delineate the frequency of admitting students who demonstrated difficulty with behaviors of competency, the same types of concerns or

difficulties that are most frequently identified and are also the ramifications of poor selection.

The frequency of selecting students who have gone on to exhibit difficulty with behaviors of competency varies in the literature. Gaubatz and Vera (2002) report that faculty in both accredited and unaccredited counseling psychology programs report having to intervene with 5.7% of enrolled students due to deficiencies. After reviewing the research, these authors also report, "between 4% and 5% of counselors in training may lack the interpersonal competence or psychological health to effectively work with clients, these figures only reflect those students that faculty members actively screened" (Gaubatz & Vera, 2002, p. 294). Forrest, Elman, Gizara & Vach-Haase (1999) reported that every year approximately 4% to 5% of counseling students are subject to remediation or dismissal. In terms of retention, the most recent calculation of annual rates for full-time graduate students in psychology were reported by Norcross, Kohout and Wicherski (2005):

> We calculated the annual retention rate for full-time graduate students in psychology as the percentage of students withdrawn or dismissed divided by the total number of full time students enrolled. The retention rate in 2003-2004 was 97.5% (attrition rate = 2.5%) for doctoral departments; 674 of 26,739 full-time students withdrew or were dismissed. The corresponding retention rate was 94.8% (attrition rate = 5.2%) for master's departments; 615 of 11,831 full-time students withdrew or were dismissed (pp. 970-971).

It is notable that these data are an improvement over previously reported findings because data from all full-time graduate students were included, whereas in the past only first-year students were reported.

In a study by Olkin and Gaughen (1991) of 100 randomly sampled, clinically oriented master's level counseling programs, the authors were not only able to provide

prevalence rates of "problem students," but also were able to identify the most common reasons the students were identified as problematic. A "problem student" was defined as someone whose difficulties were significant enough to be identified by faculty, and that faculty had to intervene in some capacity. Program directors identified how many students per year were problematic:

> Most programs (76%) identified one to three problem students per year, although almost one quarter of the programs (24%) identified four or more problem students. Only 2% reported no problem students. The mean percentage of problem students was 4.8%; the median was 3.3% (p.6).

The study provided information about how problem students were identified and the most commonly identified concerns. Students with problems were most frequently identified through academic coursework (65%), practicum (54%), faculty referral (36%) and routine evaluations (28%) (Olkin & Gaughen, 1991). Out of seven possible choices, respondents' rank ordered the top four problems of identified students. The most frequently cited problems were academic deficiencies (88%), clinical skills (77%), pervasive interpersonal problems (70%), problems related to supervision (e.g., closed to feedback, lack of self-reflection) (58%), intrapersonal problems (54%), ethical or professional misconduct (approximately 25%) and physical problems (10%). Thus, the findings suggested that it would be worthwhile to take personality factors into consideration during pre-admissions screening.

Boxley, Drew and Rangel (1986) conducted a similar study with doctoral training programs in clinical psychology and found somewhat different results. They surveyed APA-accredited internship sites to find out how many of these sites identified impaired trainees over the past five years. Respondents indicated an annual rate of 4.6% impaired trainees, but 66% said they had identified an impaired trainee in the last five years. What

differed in this survey was the most frequently reported reasons for trainee impairment. In fact, the respondents reported that personality disorders (35%) were most common, followed by depression (31%), emotional problems (31%), marital difficulties (27%) and physical illness (27%). Although academic deficiencies and substance abuse were included in the study, not many sites distinguished these circumstances as being significant reasons for trainee impairment. Given that internship sites do not necessarily deal with academics, this finding appeared to be significant because it also supported the notion that once the applied aspects of professional programs become the focus of training, psychological and personality variables become more salient factors for predicting success.

More recently, Vacha-Haase, Davenport and Kerewsky (2004) surveyed all training directors from APA-accredited clinical, school and counseling psychology programs. Of the 106 programs that responded, the following were endorsed as the most commonly occurring problems: "inadequate clinical skills (65%), defensiveness in supervision (52%), and deficient interpersonal skills (42%)" (p. 117). It is notable that 52% of the programs reported having to terminate at least one student over the course of three years, and "inadequate clinical skills" was identified as the most common reason for termination.

What ultimately happens when a student is identified as experiencing difficulties with behaviors of competency? Professional, accredited programs in counseling psychology are ethically, and in some cases legally, required to provide remediation for students who are identified with "professional performance deficiencies." However, there is no standardization of remediation or dismissal procedures (McAdams & Foster, 2007).

In response, counselor educators have developed formal remediation and dismissal policies (Baldo, Softas-Nall & Shaw, 1997; Frame & Stevens-Smith, 1995; Lumadue & Duffey, 1999). Although such procedures can effectively address the screening of students after they have been admitted to programs, the literature clearly indicates that valid variables need to be identified that can predict a candidates ability to be both academically successful and a competent practitioner.

Program directors in clinically-oriented master's level counseling programs were surveyed to identify the most common remediation methods (Olkin & Gaughen, 1991). The directors endorsed personal therapy (77%), repeating course work (70%), repeating practicum (64%) and taking a leave from the program (62%). The researchers explained that most of the students who identified as struggling with behaviors of competency are "counseled out" or simply leave the program on their own volition, compared to those who were expelled or followed through with remediation. Additionally, many students "are sent" to therapy. In this regard, it is notable that Lamb, Cochran and Jackson (1987) stated that professional counseling programs need to create better solutions in terms of remediation procedures because repeating course work or attending therapy has not been empirically validated as an effective method to improve deficient clinical skills.

There are also potential consequences for professional counseling programs when a student exhibits serious or chronic difficulties with behaviors of competency. Edwards and Schleicher (2004) identified some of the consequences:

> Selection of graduate students who perform below standards is an inefficient use of resources and may result in a weakened reputation for the university and frustration on the part of the supervising faculty members and other students. Likewise, failure to select qualified applicants is also problematic, resulting in a lost opportunity to capitalize on talent (p. 592).

In less common and more extreme circumstances, the dismissal of a student from a professional counseling program poses the risk of the student filing a lawsuit against the program and the university. *Plaintiff v. Rector and Board of Visitors of The College of William and Mary*, 2005; *Julea Ward v. Roy Wilbanks et al.,* 2010; and *Keeton v. Anderson* (Wiley, 2010) are all recent examples of lawsuits filed by graduate students against professional counseling programs in response to remediation and dismissal procedures (McAdams & Foster, 2007; Miller, 2010).

The next chapter will delineate the manner in which participants were selected. The procedures for collecting the data will then be reviewed, and a discussion of the statistical data analysis will be provided. Rationale for the inclusion and exclusion of variables will also be addressed.

CHAPTER III

METHODOLOGY

This chapter will delineate who the participants were in the study and the rationale for their selection. Additionally, a description of the data collection procedures will be provided as well as a basis for the inclusion of specific variables. Lastly, information about the instruments will be given and the method of statistical analysis will be reviewed.

Participants

The total sample included 150 individuals who graduated from a Professional Counseling Masters Program (PCMP) at a medium, public university in the Rocky Mountain region. The participants represented six consecutive entering classes who graduated between 2003 and 2009. The program was accredited by the Council for the Accreditation of Counseling and Related Educational Programs (CACREP). The participants had been selected for admissions into the program based on the following criteria: personal statement, GRE or GPA scores, three letters of recommendation, MMPI-2 assessment data, and ratings from a Pre-Admissions Workshop. It is notable that each applicant voluntarily completed the MMPI-2 and was informed that this information would be used, in part, to assist in admission decisions. The MMPI-2 was administered by a licensed psychologist and the raw data were sent to Pearson Assessments for scoring using norms for mental health and medical school applicants.

Participants, who met the requisite criteria, were admitted to the PCMP and completed an internship at the end of their training. The minimum requirements for all internships included the following: minimum of 300 direct client contact hours; minimum 600 on-site hours and minimum of one hour per week of individual supervision throughout duration of internship to be conducted by the participant's on-site supervisor. All participants who graduated from the program obtained a standardized evaluation from their internship-site supervisor which was completed and returned to the University clinical director at the end of the participant's internship.

The rationale for the selection of participants was to utilize the most current archival data available. Therefore, only participants who graduated and had complete files were included. Selection began with the most current cohort to graduate and continued in a reverse chronological process until the desired sample size, 150, was obtained. It is notable that efforts were made to use a full years worth of data prior to reversing chronological order. In other words, all of the complete files for an entire academic year were included in the sample prior to extracting data from preceding years.

Procedures

The data were archival data spanning from years 2003-2009. In order to protect the confidentiality and anonymity of participants, official university staff extracted all data from participant's files. This data were kept on file by the graduate program in compliance with university policy to retain records for all admitted students. Documented permission from the Internal Review Board at the university was obtained prior to collecting the data and performing the analysis. The following data were extracted from the participants' respective academic files: demographic information,

undergraduate GPA, mean rating score from Pre-Admissions Workshop, MMPI-2 scores on the 13 validity and clinical scales, rating scores on internship evaluations and final GPA. In order to maximize confidentiality, university staff assigned each participant a three digit ID and all data for that participant was tracked with that ID. Therefore, the data were de-identified before being received by the primary investigator. Given that the primary investigator could not ascertain the identities of those whose data were being used; there was no risk to any of the participants.

Variables

All predictor variables were components of data assessed as part of the pre-admissions process at this PCMP. Given that applicants were not required to take the Graduate Record Examination (GRE) unless their undergraduate grade point average (UGPA) was below 3.0, UGPA was used as the cognitive predictor variable. Students were also required to take the MMPI-2 and participate in a Pre-Admissions Workshop. These scores were used as the non cognitive predictor variables. Specifically, three validity scale scores and ten clinical scale scores on the MMPI-2 as well as the mean score from faculty ratings at the Pre-Admissions Workshop. The rationale for the selection of these variables was to analyze whether the prediction of academic performance could be incremented by adding non cognitive predictor variables. The rationale for using two non-cognitive predictor variables was that it allowed for the examination of whether an objective, quantitative method of screening applicants or a subjective, qualitative assessment made by the admissions committee had more incremental validity. Overall, analyzing these variables determined the utility of current admission procedures at a PCMP.

The criterion variable in the first analysis was graduate grade point average (GGPA). It was selected because this is the most commonly used criterion in predictive validity research in academia (Kuncel, Crede, & Thomas, 2005) which could aid in the generalizability of the findings. Internship evaluation mean scores were used as the criterion in the second analysis in order to test whether a criterion more in line with the PCMP training objectives could strengthen the predictor-criterion relationship in the statistical model. Additionally, using internship evaluations incorporated research. For example, Daehnert and Carter (1987) concluded that internship evaluations were the "ideal criteria" for measuring success in a PCMP and research by Leverett-Main (2004) indicated that program directors in CACREP accredited PCMP endorsed practicum/internship evaluations as the best measure of success in the program. The internship evaluations measure seven domains of professional competency on a five point Likert scale. The overall mean score was computed and then analyzed as the data from those evaluations.

Instruments

The Minnesota Multiphasic Personality Inventory-2 (MMPI-2)

The MMPI-2 is a self report inventory and a broadband measure of personality characteristics (Weiner & Greene, 2008) consisting of 567 true or false statements. It takes between 60-90 minutes to complete, requires a 6^{th} to 8^{th} grade reading level and is appropriate for individuals eighteen and older. In the basic profiles, the MMPI-2 contains nine validity scales and ten clinical scales. However, it is estimated that there are more than 120 scales that can be scored and interpreted.

Starke Hathaway, PhD, and J. Charnley McKinley, MD, published the original MMPI in 1943 (Graham, 2006). Weiner and Greene (2008) report that Hathaway and McKinley developed the original MMPI in 1940 in an effort to overcome the limitations of other instruments designed to measure personality constructs. The test authors collected data on psychopathology from various sources and then constructed items reflecting these characteristics. These items were presented to various criterion groups in the form of declarative statements to be answered true or false. The quantitative scales were then constructed based on whether the criterion group's answers differed from the normal group. "An item was tentatively selected for a scale if the difference in frequency of response between the criterion group and the normative group was at least twice the standard error of the proportions of true/false responses of the two groups being compared" (Weiner & Greene, 2008, p. 137). For example, items were linked to the schizophrenia scale if the schizophrenic sample answered the items in a statistically significant, different manner than the normal group. Accordingly, one's raw score on a specific MMPI scale represents the number of items answered the same as the criterion group. Since the items were empirically keyed, items are not necessarily descriptors of certain personality types which improves the face validity of the instrument as it often makes it difficult for respondents to differentiate what items correspond to different criterion groups.

Graham (2006) maintained that the MMPI was extremely popular as evidenced by national surveys reporting that it was the most commonly used personality inventory in the United States. However, certain aspects of the 1943 original began to be scrutinized which prompted a revision in 1989. The following lists some concerns about the original

instrument: inadequate representative standardization sample (e.g., predominantly from the Minnesota region, Caucasian, etc.); some problematic item content (i.e., language outdated, punctuation, idioms, etc.), and narrow item pool that was not broad enough to assess more contemporary concerns such as drug abuse and suicide. Accordingly, the MMPI-2 was published in 1989.

Graham (2006) cited research indicating that the MMPI-2 is the most frequently used personality test utilized by clinical psychologists. It has been included in over 2800 journal articles with many positive reviews. In fact, Nichols (1992) stated "the psycho-diagnostician selecting a structured inventory for the first time will find no competing assessment device for abnormal psychology has stronger credentials for clinical description and prediction"(p.565). Thus a more specific delineation of the psychometric properties will ensue.

The normative sample for the MMPI-2 consisted of 2600 subjects comprised of 1138 men and 1462 women. Within this sample, 841 of the individuals were married. The racial composition was 81% Caucasian, 12% African American, 3% Hispanic, 3% Native American, and 1% Asian American. The age range was 18 years to 85 years, with a mean of 41 and a standard deviation of 15.3 (Graham, 2006).

Test- retest reliability data were obtained during the revision process (Butcher, Dahlstrom, Graham, Tellegen & Kraemmer, 1989). The normative sample consisted of 82 men and 111 women who were retested after one week. The coefficients for men ranged from .67 - .93, with the Paranoid scale and the Social Introversion scale being the lowest and highest respectively. Likewise, the coefficients for women ranged from .54 to .92, with the Schizophrenia and Social Introversion scales respectively being the lowest

and highest. In general, short-term temporal stability has been shown to be as high as or higher than the original MMPI and compares favorably with other psychological tests (Graham, 2006).

It is noteworthy that the original MMPI did not focus on internal consistency reliability as it was an empirically keyed instrument. Accordingly, this is reflected on the MMPI-2 and provides an explanation as to why other logically keyed personality inventories provide more internal consistency reliability. Another consideration is that the MMPI-2 clinical scales are not uni-dimensional.

Concurrent validity data is based on MMPI and MMPI-2 raw scores. Graham (1988) reported that raw scores on the standard validity and clinical scales are greater than .98 for non-clinical and psychiatric patients. It is noteworthy that there is a discrepancy between the magnitude of raw scores as raw scores tend to be higher on the MMPI-2. Whereas subjects on the original MMPI were instructed and permitted to omit answering items, this was discouraged on the MMPI-2. Therefore, it is posited that it is not necessarily that the contemporary normative sample is more pathological but that this sample endorsed more items.

Convergent and discriminant validity data were based on a sample that consisted of 822 couples who had participated in the standardization project (Graham, 2006). The study correlated the basic clinical scales with behavior ratings from the respective partner. Results indicated that the instrument has both adequate convergent and discriminant validity.

Internship Student Evaluation

The faculty of the professional counseling masters program adopted a standardized evaluation form to be scored at the completion of the graduate student's 600 hour internship. The purpose of the feedback was to provide the clinical director with essential information to evaluate the student's effectiveness. The Internship Student Evaluation assessed the following areas of counselor competency: Opening and Rapport, Interaction and Interview Skills, Counselor Responses, Counseling Relationship, Client Conceptualization, Termination, and Case Conceptualization and Supervision. There were a total of 22 ratings on the evaluation which are scored on a 5 point Likert-type scale. There was also the option of scoring N/A if the behavior was not observed by the on-site supervisor (see Appendix C).

Sample Size

In order to estimate the optimal number of participants needed to conduct a hierarchical multiple regression analysis, the sample size was determined using an equation by Green (1991). This equation was based on a power analytic approach where alpha was set at .05 and power was set at .80 for a medium effect size (R^2 = .13 or f^2=.15). Green's equation was as follows: $N \geq L/f^2$ where $f^2 = R^2/(1-R^2)$. Based on the 15 predictor variables proposed in this study (1 for GPA; 13 validity and clinical scales on the MMPI-2 and 1 mean score on Pre-Admissions Workshop), the optimal sample size was estimated to be 139. An attempt to oversample was made so that the sample size was at least 150 participants.

Research Hypothesis and Data Analysis

H1 The prediction of graduate grade point average in a professional counseling masters program using undergraduate grade point average and

Pre-Admission Workshop rating scores can be incremented by adding personality traits, as measured by the MMPI-2.

This hypothesis was tested using hierarchical multiple regression analysis. The entry of the predictor variables was done in blocks, respectively as follows: undergraduate GPA, mean score from Pre-Admissions Workshop ratings and T-scores on three validity scales and ten clinical scales of the MMPI-2.The criterion variable was graduate GPA.

H2 The prediction of internship evaluation ratings in a professional counseling masters program using undergraduate grade point average and Pre-Admissions Workshop rating scores can be incremented by adding personality traits, as measured by the MMPI-2.

This hypothesis was tested using hierarchical multiple regression analysis. The entry of the predictor variables was done in blocks, respectively as follows: undergraduate GPA, mean score from Pre-Admission Workshop and T scores on three validity scales and ten clinical scales of the MMPI-2. The criterion variable was the mean score from the internship evaluation.

H3 Personality traits, as measured by the MMPI-2, will have more predictive validity when internship evaluations rather than graduate grade point averages are used as the criterion.

This hypothesis was tested by comparing the total percentage of variance explained by personality traits in each statistical model.

Hierarchical Multiple Regression Analysis: Assumptions and Limitations

Multiple regression analysis is appropriate when the dependent variable is metric (i.e., continuous), the independent variables are metric or dichotomous and the sample size is sufficient to ensure statistical power and generalizability (Hair, Anderson, Tatham & Black, 1995). In order to run the analysis, assumptions of normality, homoscedasticity

and linearity must be met by all individual variables. Additionally, outliers need to be examined in order to decide whether to include in the analysis or revise the model by omitting the outliers. It is not necessarily a limitation if there are violations to these assumptions because it is possible to transform the data using methods such as logarithmic, square root or inverse functions. It is important to underscore that if one of these conventions are used, it should be clearly noted in the interpretation of results (Hair et al., 1995). Before the multiple regressions were run in this study, all of these assumptions were tested.

Given that an ideal statistical model for prediction will have independent variables that are highly correlated with the dependent variable and independent variables that are not correlated with one another, it is important to check for multicollinearity. Multicollinearity is the result of high intercorrelation between different independent variables and is problematic because it will raise the standard error of coefficients and create unlikely beta estimates. Therefore, a correlation matrix should be examined and if multicollinearity is found; it will need to be fixed. There are several options to remedy multicollinearity such as omitting one of the highly correlated predictor variables or combining two of the highly correlated predictor variables (Hair et al., 1995).

Of the various procedures for conducting a regression analysis, hierarchical multiple regression analysis was utilized because it was the most appropriate method to test the hypotheses in this study. Petrocelli (2003) stated, “in hierarchical regression, the focus is on the change in predictability associated with predictor variables entered later in the analysis over and above that contributed by predictor variables entered earlier in the analysis”(p. 11). He also identified “four basic errors” in research using hierarchical

multiple regression. “They are (a) neglect of a theoretical basis for use of hierarchical multiple regression, (b) violation of causal priority, (c) use of hierarchical regression in an exploratory manner, and (d) interpretation of hierarchical regression results” (Petrocelli, 2003, p. 12). For the purpose of this discussion, these errors are considered to be limitations of hierarchical multiple regression analysis commonly overlooked in social sciences research.

The common errors were considered and there are several reasons why they did not delimit the current analyses. First, Petrocelli (2003) explained, “researchers need to generate a clear and logical rationale for its use, the selection of predictor variables, and their specific order of entry” (p.13). The theoretical basis for this study was based on empirical research that has validated that the prediction of academic performance using cognitive variables can be incremented by non-cognitive variables. The hierarchical regression tested whether the prediction of academic performance in a PCMP using standard criteria as the cognitive and first non cognitive variable can be incremented by an objective measure of personality characteristics. Therefore, the order of entry was based on empirically determining the utility of including an objective personality assessment with more commonly used admissions criteria. The assessment of utility also contributed to the determination of the specific order of entry because the least commonly used admissions criteria, objective personality assessments (MMPI-2), were the last independent variables entered into the model. Second, a violation in causal priority refers to the order of entry of the independent variables. “First, predictor variable entry should respect presumed causal priority (the direction of the causal flow). In other words, if there are causal relationships among the predictor variables the causes should be

entered into the analysis before their effects" (Petrocelli, 2003, p.14). For example, a psychological diagnosis should be entered before number of sessions attended when predicting therapeutic outcome. The causal priority for using cognitive variables prior to personality variables in the present study could be debated if one believes that personality factors have a bigger impact than intelligence on grades. Given that there is a lack of consensus on this topic in the literature, causal priority was considered a caveat when interpreting results. Third, it is considered inappropriate to use hierarchical regression in an exploratory manner because it should be a "theory driven analysis." The present study was based on empirical research evidencing that cognitive variables as well as personality characteristics could predict future academic performance. Lastly, researchers who do not understand the parameters of hierarchical regression analysis or how it differs from simultaneous or stepwise regression have reported inaccurate interpretations of the results (Petrocelli, 2003). Given that none of the shared variance was accounted for by the MMPI-2 in the present analysis, it was considered to be an important caveat in the interpretation of results.

CHAPTER IV

RESULTS

The results of this study are presented in this chapter. First, demographic data are provided to describe the sample and descriptive statistics for all study variables. Given that demographic data were gathered from the information participants provided on the graduate school form, completed as part of admission's process, it was limited to include gender, age and ethnicity. Second, correlations between the variables will be presented. Third, results of the hierarchical regression analysis are reported. Lastly, a brief discussion of how the statistical results impact each hypothesis will be provided.

Preliminary Analysis

A total of 142 participants were included in the study. In order to obtain a sample size large enough to run a hierarchical multiple regression analysis, data were extracted on graduates from a professional counseling masters program between the years of 2003-2009. It is notable that the files for a graduating class were reviewed in entirety prior to moving to the next year in reverse chronological order. Only files with complete data (study variables) for the hierarchical regression analyses were included in this study.

The gender of the participants was 15.5% male and 84.5% female. At the time of graduation, participants' ranged in age from 25 to 60 years-old and the mean age was 36.4 years-old. Only 4.6% of the participants did not report their age in the admission's packet; however, 25.4% did not report their ethnicity. Of those who did endorse ethnicity,

62% identified as White/Caucasian, 2.8% identified as Black/African American, 7.0% identified as Hispanic/Latino and 2.8% identified as Asian/Pacific Islander. A summary of the demographic information is provided in Table 1.

Table 1

Demographics Summary

Variable		Totals N	Totals %
Gender			
	Male	22	15.5%
	Female	120	84.5%
Age*	20-29	59	43.4%
	30-39	30	22%
	40-49	24	17.7%
	50-59	22	16.2%
	60-69	1	0.7%
Ethnicity**			
	White/Caucasian	88	62%
	Black/African American	4	2.8%
	Hispanic/Latino	10	7.0%
	Asian/Pacific Islander	4	2.8%

*Missing age of 6 participants
**Missing ethnicity of 35 participants

The means and standard deviations for Graduate Grade-Point Average (GGPA), Internship Evaluation Scores (IES), Undergraduate Grade-Point Average (UGPA), Pre-Admissions Workshop Rating score (PWR) and MMPI-2 scores are reported in Table 2. It is notable that all mean T scores on the MMPI-2 of the participants were within one standard deviation of the normative sample mean scores of male medical/psychology applicants.

Table 2

Descriptive statistics for Graduate Grade Point Average, Internship Evaluation Score, Undergraduate Grade Point Average, Pre-Admissions Workshop Rating and MMPI-2 Variables

Variable	M	SD	Minimum	Maximum
GGPA	3.87	.17	3.27	4.0
IES	4.53	.46	2.57	5.0
UGPA	3.45	.39	2.10	4.0
PWR	3.44	.54	1.00	4.0
MMPI-2				
L	53.91	8.87	38	76
F	43.82	6.05	36	82
K	61.23	7.37	41	76
Hs (1)	48.96	6.15	38	69
D (2)	44.08	5.56	30	64
Hy (3)	51.87	6.92	38	69
Pd (4)	53.61	8.51	37	84
Mf (5)	53.49	9.97	30	74
Pa (6)	50.92	6.98	34	70
Pt (7)	47.92	5.79	32	64
Sc (8)	48.79	6.41	31	73
Ma (9)	49.96	8.29	35	79
Si (10)	40.53	6.46	30	61

Note. N = 142

Hypotheses

Bivariate Pearson correlation coefficients for all variables in the analysis testing hypothesis one are presented in Table 3. As shown, GGPA scores were significantly correlated with UGPA and Scales 2, 5, 8 and 9. This correlation matrix was also examined for multicollinearity. Given that there were no two independent variables with a bivariate correlation of .70 or higher, all independent variables were retained in the model.

Table 3

Correlations among Undergraduate Grade Point Average, Pre-Admissions Workshop Rating, MMPI-2 and Graduate Grade Point Average (Hypothesis 1)

	GGPA	UGPA	PWR	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si
GGPA																
UGPA	.26*															
PWR	-.06	-.09														
L	-.52	-.12	-.13													
F	-.12	.01	.19*	-.19*												
K	.07	.09	-.02	.44*	-.15											
Hs (1)	-.11	.03	.17*	.10	.25*	.39*										
D (2)	-.19*	-.04	.14	-.06	.34*	-.08	.49*									
Hy (3)	.005	.08	.12	.12	.11	.26*	.69*	.35*								
Pd (4)	-.02	.10	.11	-.06	.44*	.21*	.39*	.35*	.40*							
Mf (5)	-.21*	-.07	-.14	.13	.00	.19*	.13	.05	-.03	.02						
Pa (6)	.05	.12	.13	-.08	.23*	-.06	.19	.16	.13	.24*	-.13					
Pt (7)	-.01	.08	.03	-.05	.34*	.31*	.43	.40*	.36*	.48*	.10	.29				
Sc (8)	-.17*	.07	.15	-.00	.50*	.39*	.57	.29*	.38*	.60*	.17	.26	.64			
Ma (9)	-.28*	-.70	-.07	-.18*	.29*	-.32*	-.04	-.02	-.04	.20*	.01	.13	.13	.31*		
Si (10)	-.04	-.05	.10	-.05	.25*	-.30*	.09	.30*	-.15	-.02	-.04	-.00	.14	-.03	-.22*	

*indicates significance of .05 or less

Table 4 shows bivariate Pearson correlation coefficients for all variables in the analysis testing hypothesis two. As shown, only Scale 2 significantly correlated with internship evaluation ratings. This correlation matrix was also examined for multicollinearity. Given that there were no two independent variables with a bivariate correlation of .7 or higher, all independent variables were retained in the model (Pallant, 2007).

Table 4

Correlations among Undergraduate Grade Point Average, Pre-Admissions Workshop Rating, MMPI-2 and Internship Evaluation Rating (Hypothesis 2)

Variable	IER	UGPA	PWR	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si
IER																
UGPA	.10															
PWR	.05	-.09														
L	.07	-.12	-.13													
F	.09	.01	.19*	-.19*												
K	.07	.09	-.02	.44*	-.15											
Hs (1)	.02	.03	.17*	.10	.25*	.39*										
D (2)	-.21*	-.04	.14	-.06	.34*	-.08	.48*									
Hy(3)	.01	.08	.12	.12	.11	.26*	.69*	.35*								
Pd (4)	.02	.10	.11	-.06	.44*	.21*	.39*	.35*	.40*							
Mf (5)	-.01	-.07	-.14	.13	.00	.19*	.13	.05	-.03	.02						
Pa (6)	.12	.12	-.13	-.08	.23*	-.06	.19*	.16	.13	.24*	-.13					
Pt (7)	-.08	.08	.03	-.05	.34*	.31*	.43*	.40*	.36*	.48	.10	.29				
Sc (8)	-.03	.07	.15	-.002	.50*	.39*	.57*	.29*	.38*	.60	.17*	.26*	.64*			
Ma(9)	-.12	-.07	-.08	-.18*	.29*	-.32*	-.04	-.02	-.04	.20*	.01	.13	.13	.31*		
Si (10)	-.06	-.05	.10	-.05	.25*	-.30*	.09	.30*	-.15	-.02	-.04	-.003	.14	-.03	-.22*	

*indicates significance at .05 or less

Hypothesis 1

The first hypothesis of the study stated:

> H1 The prediction of graduate grade-point average in a professional counseling masters program using undergraduate grade-point average and Pre-Admissions Workshop rating scores can be incremented by adding personality traits, as measured by the MMPI-2.

First, all variables were analyzed to ensure the assumptions of hierarchical multiple regression were met; namely, linearity, constant variance (homoscedasticity) and normality. Scatter plots and residual plots did not indicate any nonlinear relationships between the independent and dependent variables or violations of constant variance. Tests for normality indicated that there were mild violations of normality for the PWR and GGPA variables.

According to Osborne (2010),

> Some research has shown that parametric tests (e.g., multiple regression, ANOVA) can be robust to modest violation of these assumption. Yet the reality is that almost all analyses (even nonparametric tests) benefit from improved normality of variables, particularly where substantial non-normality is present (p.1).

Similarly, Hair et al. (1995) indicated,

> While regression analysis has been shown to be quite robust even when the normality assumption is violated, analysts should estimate the regression analysis with both the original and transformed variables to assess consequences of nonnormality of the independent variables on the interpretation of results (p.135).

Accordingly, square root transformations were performed because Hair et al. (1995) indicated this procedure is best when there is a negatively skewed distribution as was the case for both of the variables Pre–Admissions Workshop ratings (PWR) and graduate grade point average (GGPA). This transformation for PWR produced a Skewness value of .66 and a Kurtosis value of .60, which is in the range of a normal distribution (Hair et al, 1995). This transformation did not improve normality for the GGPA distribution;

therefore, several other transformations were performed including, logarithmic transformations and Box-Cox transformation (Osborne, 2010). The transformation that produced the most normal distribution entailed reflecting the distribution and then applying an inverse transformation. The skewness and kurtosis of the distribution improved; however, the skewness value was still slightly out of range (-1.03). The impact of the transformed variables on the results was then assessed by running the analysis with the original and transformed data. The analysis with the transformed values produced almost the same results and the same three variables were found to significantly contribute to the model. Given that the fit of the model was not significantly improved after performing transformations and the consequences of nonnormality were assessed as being minimal, the original data were retained for the final analysis.

In addition to the examination of the correlation matrix, "collinearity diagnostics" were also produced to check for multicollinearity. Results of both the collinearity tolerance (all values greater than .10) and variance inflation factors (VIF) (all values less than 10.0) suggest that the estimated β are well established in the model. Scatter plots were also examined for outliers. The standardized residual plot indicated there was one data point with a value higher than -3.3, which is the cutoff that Tabachnick and Fidell (2007, p.128) use to identify outliers. However, it was retained in the analysis because it was assessed that it represented a valid observation in the population and was not due to data entry error or an extraordinary event (Hair, et al., 1995). Lastly, the Durbin-Watson test statistic (1.83) indicated that the residuals are independent and the assumption of independent errors was met.

Hierarchical multiple regression analysis was used to determine the incremental validity of the MMPI-2 in the prediction of graduate GPA (GGPA) in a professional counseling masters program after controlling for more commonly used criteria, undergraduate GPA and Pre-Admissions Workshop rating (PWR) scores. Undergraduate GPA (UGPA) was entered at the first block, explaining 6.6% of the variance in GGPA, F (1,140)=9.97, p=.002 (R^2 = .066, ΔR^2= .066). After entering PWR scores in block two, the total variance explained by the model as a whole was 6.7%, F(2,139)= 4.99, p=.008 (R^2 = .067, ΔR^2= .001). When the MMPI-2 scores were added in block three, the final model accounted for 24.4% of the variance, F (15,126)=2.80, p=.001 (R^2 = .244, ΔR^2= .177). See Table 5 for the hierarchical regression analysis results. In the final model, three variables were statistically significant. UGPA had the lowest beta value (*beta*=.188, p=.021), then Scale 2 on the MMPI-2 (*beta*= -.237, p=.025) and then Scale 9 (*beta* = -.306, p=.009). Given that the MMPI-2 contributed an additional 17.7% of the variance above and beyond what was predicted by UGPA and Pre-Admissions Workshop ratings, Hypothesis 1 was supported.

Table 5

Hierarchical regression analysis of Graduate Grade Point Average on predictor variables (Hypothesis 1)

Predictor Variable	*B*	*SE*	*Beta*	*t*	*p* value
Step 1 (R^2 = .066, ΔR^2= .066)					
Constant	3.49	.121		28.76	.000
UGPA	.110	.035	.26	3.16	.002
Step 2 (R^2 = .067, ΔR^2= .001)					
Constant	3.47	.14		24.14	.000
UGPA	.11	.04	.26	3.11	.002
PWR	.007	.03	.02	.29	.770
Step 3 (R^2 = .244, ΔR^2= .177)					
Constant	4.28	.35		12.24	.000
UGPA	.08	.03	.19	2.34	.021*
PWR	.006	.03	.02	.23	.820
L	-.001	.002	-.06	-.59	.557
F	.000	.003	.02	.16	.876
K	-.001	.003	.03	-.21	.838
1	-.001	.004	-.02	-..13	.894
2	-.007	.003	-.24	-2.26	.025*
3	.001	.003	.03	.19	.848
4	.002	.002	.10	.92	.358
5	-.003	.001	-.15	-1.81	.072
6	.001	.002	.05	.53	.598
7	.005	.003	.19	1.55	.124
8	-.005	.004	-.18	-1.18	.241
9	-.006	.002	-.31	-2.66	.009*
10	-.002	.003	-.07	-.64	.523

Note. N = 142.
*indicates significance of .05 or less

Hypothesis 2

The second hypothesis of the study stated:

> H2 The prediction of internship evaluation ratings in a professional counseling masters program using undergraduate grade point average and Pre-Admissions Workshop rating scores can be incremented by adding personality traits, as measured by the MMPI-2.

Prior to estimating the regression equation, the assumptions for regression analysis were tested. Scatter plots and statistical tests evidenced support that there were no violations of the assumptions of linearity and constant variance; however, there were

mild violations of the normality assumption for the PWR and internship evaluation rating (IER) variables. Although regression analysis has been shown to be robust to mild normality violations (Hair et al., 1995), data transformations were performed to assess whether the normality of the distributions could be improved and the model strengthened. A square-root transformation was performed on the PWR data. This resulted in the skewness and kurtosis of that distribution to fall within the range indicative of a normal distribution. The transformation of the internship evaluation ratings data first consisted of reflection of the data and then a logarithmic transformation was indicated. This data transformation satisfied the criteria for a normal distribution as the skewness (.530) and kurtosis (-.69) of the distribution fell within the range of a normal distribution. It is notable that all data transformations were computed and examined prior to determining which method best achieved the desired results.

Examination of the correlation matrix indicated that all bivariate correlations were equal to or less than .70, which suggested that multicollinearity was not a concern (Pallant, 2007). Collinearity diagnostics were also performed to check for multicollinearity, collinearity tolerance (all values greater than .10), and variance inflation factors (VIF)(all values less than 10.0). All values were within range which suggested that the estimated β were well established in the model (Pallant, 2007) and violations of collinearity were not committed.

Scatter plots were then examined for outliers. One multivariate outlier was identified after checking Mahalanobis distances; however, the case was retained in the analysis because it was assessed to represent a valid observation in the sample and was not due to data entry error or an extraordinary event (Hair, et al., 1995). Lastly, the

Durbin-Watson test statistic (2.02) indicated that the residuals are independent and the assumption of independent errors was met.

First, the original variables were used in the hierarchical multiple regression analysis to determine the incremental validity of the MMPI-2 in the prediction of performance on internship evaluations in a professional counseling masters program after controlling for more commonly used criteria, undergraduate GPA and Pre-Admissions Workshop rating scores. Undergraduate GPA was entered at the first step, explaining .9% of the variance in internship evaluation ratings, F (1,140) = 1.28, p=.260 (R^2 = .009, ΔR^2= .009). After entering Pre-Admissions Workshop rating scores in block two, the total variance explained by the model, as a whole, was 1.1%, F(2,139)= .764, p=.468 (R^2 = .011, ΔR^2= .002). When the MMPI-2 scores were added in step three, the final model accounted for 15.6% of the variance, F (15,126) =1.554, p=.096 (R^2 = .156, ΔR^2= .145). Although the additional incremental validity of the MMPI-2 was substantial, Hypothesis 2 was rejected at the significance level of .05 (p=.096). However, it is important to underscore that calling a model significant is somewhat arbitrary in that the researcher sets the significance level. The model was significant at the .10 level and those results are presented in Table 6. It is also notable that the p-values for the Type III Sum of Squares were significant at the .05 level for MMPI-2 Scale F and Scale 2. This finding indicated that the unique contribution of each of those independent variables was significant after adjusting for the variability explained by all other independent variables.

It is notable that once the regression model had been estimated, another analysis was performed using the transformed data in order to assess the consequences of mild normality violations. Nearly identical results were produced when the original values

were used to estimate the regression model. The total model accounted for 14.6% of the variance in performance outcomes and the unique variance accounted for by the MMPI-2 was 13.6%. Likewise, none of the variables significantly contributed to the model and an equation could not be produced. Therefore, the variables transformed to meet the assumptions of multiple regression analysis accounted for slightly less variance in the model. However, the transformed variables did not significantly change the results of the analysis or ability to support Hypothesis 2.

Table 6

Hierarchical regression analysis of Internship Evaluation Score on predictor variables (Hypothesis 2)

	B	*SE*	*Beta*	*t*	*p* value
Step 1 (R^2 = .009, ΔR^2=.009)					
Constant	4.144	.347		11.955	.000
UGPA	.113	.100	.095	1.131	.260
Step 2 (R^2 = .011, ΔR^2= .002)					
Constant	4.034	.410		9.836	.000
UGPA	.108	.100	.091	1.076	.284
PWR	.037	.073	.043	.505	.614
Step 3 (R^2 = .156, ΔR^2= .145)					
Constant	4.720	1.026		4.603	.000
UGPA	.075	.101	.063	.740	.461
PWR	.020	.075	.023	.265	.792
L	.005	.005	.089	.874	.384
F	.019	.008	.254	2.306	.023*
K	-.004	.010	-.068	-.417	.677
1	.016	.012	.208	1.342	.182
2	-.029	.009	-.351	-3.163	.002*
3	-.003	.009	-.041	-.307	.759
4	.004	.006	.077	.681	.497
5	.002	.004	.033	.376	.707
6	.008	.006	.127	1.377	.171
7	-.002	.010	-.026	-.206	.837
8	-.011	.011	-.151	-.960	.339
9	-.010	.007	-.186	-1.527	.129
10	-.007	.008	-.096	-.809	.420

Note. N = 142.
*indicates significance of .05 or less

Hypothesis 3

The third hypothesis stated:

> H3 Personality traits, as measured by the MMPI-2, will have more predictive validity when internship evaluations rather than graduate grade point averages are used as the criterion.

Originally, this hypothesis was to be tested by comparing the total percentage of variance explained by personality traits in each statistical model. Given that the MMPI-2 explained 17.7% of the unique variance in GGPA and 14.5% of the unique variance in

the performance on internship evaluation ratings; Hypothesis 3 was not supported. However, the main reason this hypothesis was rejected was because contrary to the prediction of GGPA, none of the models (overall *p* values) were statistically significant predictors of performance on internship evaluations.

A more thorough narrative discussion of these results will be presented in Chapter 5. The implications of both the statistical and practical significance of the hypothesis testing will also be reviewed. Finally, the implications of the findings for the field of professional psychology will be provided and future directions for empirical research will be offered.

CHAPTER V

DISCUSSION

The purpose of this study was to investigate whether or not there is empirical support to include an objective measure of personality characteristics as part of standard admissions procedures in a professional counseling masters program (PCMP). This chapter will discuss each hypothesis in order to review the results presented in the previous chapter, highlight the implications to the field of professional psychology and provide suggestions for future research.

Undergraduate Grade Point Average, Pre-Admissions Workshop Ratings, MMPI-2 and Graduate Grade Point Average

As first hypothesized, the MMPI-2 provided incremental validity in the prediction of GGPA above what could be predicted by UGPA and PWR. In fact, results from the hierarchical multiple regression analysis indicated that the overall model accounted for 25% of the variance in GGPA and of that, the MMPI-2 independently explained 17% of the variance. UGPA was significantly associated with GGPA *(r = .26; p<.05)* and produced a statistically significant change in R^2 *(beta .188, p=.021; p<.05)* in the analysis. Consistent with the Pearson correlation coefficients, PWR was not significantly associated with GGPA and did not significantly contribute to the model.

In the first model, it was estimated that a one standard deviation increase in UGPA would yield a .19 increase in GGPA. Given the consensus in the literature that

cognitive measures can predict academic outcomes (Burton & Ramist, 2001; Furnham, Monsen & Ahmetoglu, 2009; Trapmann et al 2007) this study lends further support. Furthermore, it upholds the notion that when building a model to predict academic outcomes it is prudent to include both cognitive and non cognitive variables. For example, empirical support from meta-analyses and large scale studies have evidenced that the prediction of academic performance using cognitive variables (e.g., GPA or academic achievement tests) can be incremented by adding objective measures of personality (Lievens et al., 2009; Poropat, 2009; Trapmann et al., 2007).

Conversely, the non cognitive variable, PWR, did not significantly contribute to the prediction of GGPA *(ΔR^2=.001)*. This finding brings into question the utility of subjective personality assessment as part of admissions screening, given research indicated that subjective methods are the most common admissions method of screening for personality or character currently used by PCMPs. However, it is also important to underscore that, in general, it may not be that subjective assessments of personality characteristics are poor predictors of performance in a PCMP; rather it could have been this particular variable performed poorly. In other words, it is likely that since PWRs were based on a 4 point Likert-type rating scale and the majority of participants (i.e., admitted students) scored between three and four on the rating scale; restriction of range limited the predictive ability of this variable in this analysis.

Although Pearson correlation coefficients indicated that Scales 2 (D), 5 (Mf), 8 (Sc) and 9 (Ma) significantly correlated with GGPA ($p<.05$), regression analysis results showed Scale 9 (Hypomania) and 2 (Depression) to be the significant contributors to the prediction of GGPA. In fact, a one standard deviation decrease in score on the

Hypomania scale would yield a .306 increase in GGPA, suggesting that individuals with higher scores on the Hypomania scale tended to have lower GGPA scores. This finding is consistent with extant literature on predictor – criterion relationships using the MMPI/MMPI-2 and performance outcomes (Compton, 2009; Hiatt & Hargrave, 1988; King & Bailey, 2002). For example, the Ma scale predicted undergraduate grade point average in Compton's (2009) study, it was shown to predict "problem" officers by Hiatt and Hargrave (1988) and King and Bailey (2002) found that high scores correlated with poor academic performance as measured by undergraduate GPA.

As would be predicted from similar studies (Davis & Widseth, 1978; Strupp & Bloxom, 1975) the Depression scale (2) was a statistically significant predictor of GGPA. This result is especially noteworthy because results were replicated using the MMPI-2 rather than the original MMPI. Specifically, results from this study indicated that a one standard deviation decrease in score on the Depression scale (2) would yield a .24 increase in GGPA. According to Graham (2006), people who score high on the Depression scale "tend to feel hopeless and to be pessimistic about the future in general and more specifically about the likelihood of overcoming their problems and making better adjustment" (p. 69). Given that professional counselors aim to instill hope and assist people make better adjustment, the negative relationship seems logical. Findings also imply that students with greater levels of depression may have more difficulty focusing on academic tasks than those with fewer symptoms of depression.

Undergraduate Grade Point Average, Pre-Admissions Workshop Ratings, MMPI-2 and Internship Evaluation Ratings

As predicted, the MMPI-2 showed incremental validity, over UGPA and PWR, in the prediction of internship evaluation ratings, but not at the .05 alpha level set for the study. The total model accounted for 15.6% of the outcome variance and of that, 14.5% was explained by the MMPI-2 predictor variables. Given that the p values were not significant for any of the models at the .05 alpha level set for this study, a statistical equation could not be built. However, the final model with all three blocks was significant at the .10 alpha level, which indicated that there is a one in ten chance that a researcher will reject the null hypothesis when the null hypothesis is in fact true.

Specifically, Scale F (p=.023) and Scale 2 (p=.002*)* on the MMPI-2 each significantly contributed the prediction of IER at the .05 alpha level. Scale F is called the Infrequency Scale and high scores are typically associated with individuals with severe pathology or with a desire to appear more psychologically disturbed than they are in actuality (Graham, 2006). Given that that Scale F and IER were positively associated, it means that as the infrequency scores increase so do IER. It is notable; however, that the mean T- score (T=43; SD=6) for the participants was well within the average range. According to Graham (2006) scores in this range for nonclinical populations indicate that the respondent answered in an acceptable manner and may be endorsing some deviant beliefs. "Sometimes scores at this level indicate persons with deviant social, political, or religious convictions. Persons with scores at the upper end of this range may be accurately reporting psychological problems"(Graham, 2006, p. 29). Thus, the relative

elevations may actually make the student better able to empathize with a client's emotional struggles.

As previously mentioned, Scale 2 is the Depression scale and given that the findings are similar to those found when testing Hypothesis 1, similar conclusions can be drawn. Participants who endorsed more symptoms associated with depression at the beginning of the PCMP tended to have lower IER at the end of the PCMP. These findings suggest that due to increased symptoms of depression, participants may have had difficulty in the role of a counselor or that they may have had more difficulty focusing on academic tasks than those with fewer symptoms.

In a similar study using the MMPI and internship evaluations, Daehnert and Carter (1987) found very different results. Correlations between MMPI-2 scale scores and internship evaluations ranged from -.21 to .56 and in this study the strongest correlation was -.21. Given the sample populations were both graduate psychology students, it seems likely that different results were achieved because this study used the mean score from internship evaluations rather than scores from distinct categories. As Oswald et al. (2004) stated, "The number of dimensions should not be so many that the information is unwieldy, yet not so few that the domain of college performance is oversimplified and not appropriately represented" (p.187). It appears possible that using the mean score oversimplified the criterion and that stronger correlations may have become apparent if the multidimensional evaluation categories had been utilized.

Given that the MMPI-2 performed as predicted but not at the .05 alpha level set for the study, it is important to discuss the implications of the Hypothesis 2 results in terms of practical verses statistical significance. According to Kirk (1996), "Statistical

significance is concerned with whether a research result is due to chance or sampling variability; practical significance is concerned with whether the result is useful in the real world"(p.746). He argued that the myopic focus on null hypothesis significance testing is problematic in professional psychology research. He stated, "Unfortunately, all too often the primary focus of research is on rejecting a null hypothesis and obtaining a small *p* value. The focus should be on what the data tell us about the phenomenon under investigation" (Kirk, 2001, p.213). Therefore, the practical significance of the results for Hypothesis 2 are that the MMPI-2 can provide incremental validity and predict performance on IER in a PCMP beyond what can be accounted for by traditional admissions criteria. Additionally, it lends empirical support to previous findings that the predictive validity of personality increases at higher levels of education where the pool of applicants is more homogenous and the criterion of success involves application of knowledge to applied practice. Finally, it demonstrated that in addition to personality instruments based on the Five Factor Model (FFM), the MMPI-2 can also provide incremental validity in the prediction of academic outcomes.

MMPI-2, Grade Point Average, Internship Evaluations and Criterion Issues

It was hypothesized that a stronger statistical model and more outcome variance could be explained when the criterion of success was more in line with the training objectives at the PCMP. This hypothesis was inspired by the well documented issue in predictor – criterion and incremental validity research.

> Within the complex and competitive admissions process, colleges and universities seek out the best students possible for their institutions, in which "best" can be defined in many ways….However, critics argue there is substantial room for

> improvement with respect to the validity and practical utility of current selection tools" (Oswald et al., 2004, p.187).

In both analyses, the MMPI-2 had the most independent and incremental effects on the outcome. However, the most variance was explained when GGPA was the criterion and when IER was the criterion, none of the models were statistically significant at the .05 alpha level set for the study. Therefore, Hypothesis 3 was rejected.

Implications of the Study and for Professional Counseling Masters Programs

The results of this study contribute to the literature in several significant ways. First, it added to the sparse research on predictive validity at the graduate level and the expressed need to identify other criterion of graduate education success as well as other predictor variables of success in graduate school. Second, existing research on personality traits and academic success had used instruments based on the Five Factor Model (FFM) and were conducted for research purposes only. Therefore this study diversified that line of inquiry and evidenced that like instruments based on the FFM (Chamorro-PreMuzic & Furnham, 2003; Trapmann, Hell, Hirn & Schuler, 2007), the MMPI-2 has independent and incremental effects on academic outcomes, after traditional predictors of those outcomes are controlled. Likewise, the results demonstrated that the MMPI-2 was still able to significantly contribute to the prediction models irrespective of the inherent range restriction of scores that resulted from using a sample already prescreened with the predictor variables. Additionally, this study confirmed that MMPI-2 scores on the Depression (2) and Hypomania (9) scales are indicators of future performance (2-3 years after the MMPI was completed). Accordingly, this study has

provided empirical evidence to support the utility of including valid, reliable personality instruments in PCMP admissions procedures.

It is beyond the scope of the results to ascertain with certainty how well PWR would have predicted performance in a PCMP without the restriction of range of scores. However, it is quite likely that had the PWR ratings not been used to make admissions decisions (i.e., majority of accepted applicants had high scores), these ratings would have accounted for more of the variance than found in this study. Accordingly, results should be interpreted with the caveat that the results show a lack of predictive validity for PWR only within a restricted range of scores. Additionally, it highlights, once again, that even with the inherent range restriction that results when using participants who have been previously screened and admitted into a competitive graduate programs, the MMPI-2 performed well and demonstrated utility for admissions procedures. This finding warrants a brief exploration as to why there is merit in quantitatively assessing personality traits rather than relying on commonly used qualitative assessments. The following explanation about the shortcomings of relying on subjective assessment for college admissions was given by Oswald et al. (2004):

> Popular methods of obtaining such information include achievement test scores, letters of recommendation, personal statements, lists of extracurricular activities, interviews, and peer references. There exists some support for the incremental validity and practical usefulness of such measures over the more common predictors mentioned above. However, these supplementary measures are problematic to the extent that (a) admissions personnel pay attention to, interpret, and weight this information in different ways; (b) admissions personnel rely on information about students' past experiences that is to some extent idiosyncratic and not in a standardized format; (c) collecting and evaluating this information requires extra cost in time and resources; and (d) information is self reported and may be difficult to verify. Not implementing, scoring, or weighting such measures in a systematic manner across colleges, and not tying these measures to a relatively broad domain of college performance where supplementary measures may be more useful, preclude a solid conceptual understanding and a consistent

> and practical level of incremental validity above standardized test scores and high school GPA (p.188).

This could be of interest to applied researchers and admissions committees given that subjective assessments of personality characteristics (e.g., letters of recommendation, interview, and personal statements) are the most common way admissions committees screen (Johnson & Cambell, 2004; Norcross, Kohout& Wicherski, 2005).

There are several additional implications from this study for PCMPs. First, empirical support for the inclusion of an objective personality assessment as part of standard admissions in a PCMP has been demonstrated. In this study, the MMPI-2 was the instrument used in the predictive model; however, instruments based on the Five Factor Model have also been shown to provide incremental validity of personality traits in the prediction of academic outcomes (Poropat, 2009; Trapmann et al., 2007). Given that the MMPI-2 has 567 questions and takes between 60-90 minutes to complete (Weiner & Greene, 2008), it may be worthwhile for PCMP to test whether personality instruments based on the FFM with fewer questions, such as the NEO-PI-R (240 questions) or the NEO-FFI (60 items) (Costa & McCrae, 1991, 1992) can predict a commensurate amount of variance in performance as the MMPI-2 did in this study. This may improve the utility or practicality of incorporating the findings of this study into standard admissions procedures at other PCMPs.

Finally, PCMPs should consider the practical significance of the finding that participants who scored higher on the Depression Scale (2) and the Hypomania Scale (9) had lower performance outcome scores. It is important to underscore that relatively speaking, even those participants who scored higher on Scales 2 and 9, persevered well in the program as evidenced by graduating. Therefore, it is suggested that applicants with

higher scores on Scales 2 and 9 be identified in order to direct them towards resources that could assist them improve their mental health rather than as a method to screen out the applicant from admissions. This seems especially important given that as individuals progress in graduate programs in psychology, the overall number of distressed students increase and the level of severity of their distress increases as well (Kuyken, Peters, Power & Lavender, 2003).

Limitations and Directions for Future Research

There were several limitations to this study. Accordingly, these will be reviewed and possible remedies will be offered. First, this study was retrospective; therefore, the analysis was bound by the data that had been collected as part of admissions at the PCMP which would affect the ability to generalize results to PCMPs with different admissions criteria. Accordingly, future research could replicate this study with different cognitive as well as non cognitive predictors and the MMPI-2. For example, a multivariate analysis using GRE as the cognitive variable in the first block, letters of recommendation as the subjective assessment of personality traits in the second block and the MMPI-2 in the third block.

Second, the participants had been screened and admitted to the PCMP which likely affected the range of scores. Similarly these participants were seeking admissions to a competitive MA program and may have attempted to present themselves in a particularly favorable manner which may not have accurately reflected actual abilities or traits. Given that the participants would generally be considered high achieving, it is recommended that the study be replicated and that range restriction be statistically accounted for. However this possibility is somewhat mitigated by the fact that if an

applicant had a MMPI-2 profile that was considered to be invalid, the applicant was required to retake the instrument provided that demographic variables (e.g., applicant spoke English as a second language) were not a possible factor.

Third, the participants all attended a PCMP at a medium university in the Rocky Mountain region. Given that the findings may not generalize to other countries or regions of the United States, it is recommended that this study be replicated at PCMPs in other regions.

Fourth, there were delimitations in this study, which primarily related to the design. The decision to delimit the MMPI-2 variables to the ten basic profiles and three of the nine validity scales was guided by parameters of the multivariate analysis. Based on the significant findings of this study, future researchers could use the MMPI-2 subscales, such as the Harris-Lingoes, in an analysis to better understand how the dimensions of Depression and Hypomania relate to performance in a PCMP. Similarly, it would be interesting to go beyond the 13 scales used in this study and examine configural aspects of the MMPI-2 scale scores. For example, an examination of code types or of cases where scale1 and Scale 3 are both ten T-score points higher than Scale 2 (i.e., suggestive of somatization) (Graham, 2006) might provide a better understanding as to how personality and performance in a PCMP are related.

Fifth, it is important to address how using grade point average as variables likely affected the results. In terms of graduate grade point average as the criterion, the use of a sample of graduate students from a single academic discipline, with a standardized curriculum with which to compare grade point averages was more reliable and consistent than would be the case if different academic disciplines or programs from other

universities were part of the sample. Conversely, the use of undergraduate grade point average did not have the same benefits. As Oswald et al. (2004) explained,

> If one takes seriously what colleges claim to want in their students, then we argue that it is appropriate to reconsider traditional GPA and graduation criteria on several accounts. First, traditional academic outcomes are useful for what they are intended to measure but insufficient when one considers the entire experience contributing to students' performance and success in college. Furthermore, GPA as a composite measure is not standardized and may represent the outcome of some very different student behaviors, as reflected in different types of courses taught by different instructors (p. 187).

Therefore, this lack of standardization is a caveat because it is difficult to assess whether there are true differences in scores. Additionally, it is doubtful that the UGPA is a reflection of purely cognitive abilities. It would be beneficial if this study was replicated either substituting the cognitive variable with GRE scores or adding a block for the GRE scores to better understand if some of the variance explained by undergraduate GPA could in actuality be more related to personality. In other words, one's GPA reflects more than just intellect whereas the GRE seems to be more of a quantitative and objective assessment of cognitive abilities.

Lastly, the findings of Hypothesis 2 could be interpreted to suggest that the MMPI-2 measured characteristics that are distinct from those measured by internship evaluations. However, given the findings by Daehert and Carter (1987) it is also possible that the statistical analysis did not allow for a more comprehensive examination of the relationship between personality factors and internship skills. In other words, the mean score of the internship evaluation ratings may have been too narrow and an analysis that could examine subcategories of the IER as multiple dependent variables might provide a more accurate depiction of the correlations. For instance, what scales and T – score elevations are associated with being open to supervision/feedback or with ethical

practice? This could be achieved by computing Pearson correlation coefficients between each of the MMPI-2 scales and the sub category mean scores on the internship evaluations. Another method for better understanding whether the candidate has characteristics that match desired attributes for a becoming a professional counselor would be to have PCMP faculty rate students on their clinical skills whereby they could be divided into top, middle and bottom thirds. A MANOVA or a discriminant analysis could then analyze the data using the top and bottom third as the independent variables and the MMPI-2 as the dependent variables.

Summary

The purpose of this study was to investigate whether there is empirical support to include an objective measure of personality characteristics as part of standard admissions procedures in a professional counseling masters program (PCMP). As hypothesized, the MMPI-2, an objective personality instrument, had incremental and independent effects on two academic outcomes; namely, graduate grade point average (GGPA) and internship evaluation ratings (IER). In fact, both analyses evidenced that after controlling for traditional admissions criteria, undergraduate grade point average (UGPA) and Pre-Admissions Workshop ratings (PWR), the MMPI-2 accounted for the most variance in both outcome variables. These results imply that personality characteristics are veritably relevant to academic performance and advance prior research indicating that objective measures of personality are predictive of academic outcomes. As Poropat (2009), who conducted the most comprehensive meta-analytic investigation of personality-academic performance relationships concluded, “Personality should take a more prominent place in future theories of academic performance and not merely as an adjunct to intelligence” (p.

333). The third hypothesis sought to identify criterion other than GGPA as a measure of success and tested whether a stronger statistical model could be built when using a criterion more in line with the training objectives of a PCMP. This was rejected because more variance was explained in the model using GGPA and an equation could not be built when using IER. However, using IER as a criterion of success in a PCMP did demonstrate potential to overcome the limitations of using GPA and future research should focus on a way to capture the different dimensions of internship performance when analyzing it as a criterion. Overall, the practical significance of the present study is that it moves the field of professional psychology closer towards developing improved admissions criteria for graduate school admissions and underscores the importance of including personality characteristics when building models to predict performance outcomes.

REFERENCES

American Psychological Association Board of Educational Affairs. (2007). *Assessment of competency benchmarks work group: A developmental model for the defining and measuring competence in professional psychology.*

Baldo, T. D., Softas-Nall, B. C., & Shaw, S. F. (1997). Student review and retention in counselor education: An alternative to Frame and Stevens-Smith. *Counselor Education and Supervision, 35,* 245–253. Retrieved from http://www.counseling .org/Publications/Journals.aspx

Banks, Mooney, Mucowski, & Williams, (1984). Progress in the evaluation and prediction of successful candidates for religious careers. *Counseling and Values, 28,* 82-91. Retrieved from http://www.counseling.org/Publications/Journals.aspx

Boxley, R., Drew, C. R., & Rangel, D. M. (1986). Clinical trainee impairment in APA approved internship programs. *The Clinical Psychologist, 39,* 49–52. Retrieved from http://www.apa.org/divisions/div12/clinpj.html

Brooks, J., Holtum, S., & Lavender, A. (2002). Personality style, psychological adaptation and expectations of trainee clinical psychologists. *Clinical Psychology and Psychotherapy, 9*,253-270. doi:10.1002/cpp.318

Burton, N.W., &Ramist, L. (2001). *Predicting success in college: SAT studies of classes graduating since 1980*. The College Board Research Report, 2001–2002. New York: College Entrance Examination Board.

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, achievement motivation and academic success of psychology students in higher education. Personality and Individual Differences, *29*, 1057–1068. doi:10.1016/S0191-8869(99)00253-6

Butcher, J.N. (1979). Use of the MMPI in personnel selection. In Butcher, J.N. (Ed.), *New developments in the use of the MMPI*. Minneapolis: University of Minnesota Press.

Butcher, J.N. (1985). Personality assessment in industry. In Bernardin, H.J., & Bownas, D. (Eds.), *Personality assessment in organizations.* New York: Praeger.

Butcher, J.N. (1991). Screening for psychopathology: Industrial applications of the Minnesota Multiphasic Personality Inventory (MMPI-2). Chapter in J. Jones, B.D. Steffey, & D. Bray (Eds.). *Applying psychology in business: The manager's handbook.* Boston: Lexington Books.

Butcher, J.N. (1994). Psychological assessment of airline pilot applicants with the MMPI--2. *Journal of Personality Assessment, 62*, 31-44. doi:10.1207 /s15327752jpa6201_

Butcher, J.N. (2002). Assessing Pilots with the 'Wrong Stuff': A Call for Research on Emotional Health Factors in Commercial Aviators. *International Journal of Selection and Assessment, 10* (1/2)*,* 168 – 184. doi:10.1111/1468-2389.00204

Butcher, J.N., Dahlstrom, W.G., Graham, J.R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory-2 (MMPI-2): Manual for administration and scoring.* Minneapolis: University of Minnesota Press.

Butcher, J. N., Graham, J. R., Dahlstrom, W. G., & Bowman, E. (1990). The MMPI-2 with college students. *Journal of Personality Assessment, 54*(1), 1-15. doi:10.1207/s15327752jpa5401&2_1

Butcher, J. N., Ones, D.S. & Cullen, M. (2006). Personnel screening with the MMPI-2. In J.N. Butcher (Ed.). *MMPI-2: The practioner's handbook*. (pp. 381-406) Washington, D.C.: American Psychological Association. doi:10.1037/11287-014

Butcher, J.N. & Rouse, S.V. (1996). Personality: Individual differences and clinical assessment. *Annual Review of Psychology*, *47*, 87-111. doi:10.1146/annurev .psych.47.1.87

Callan, R.J. (1997). An attempt to use the MMPI as a predictor of failure in military training. *Military Psychology, 121,* 533-557. Retrieved from http://0-www .informaworld.com.source.unco.edu/smpp/title~content=t775653681~db=all

Carbone, E.G., Cigrang, J.A., Todd, S.L., & Fiedler, E.R. (1999). Predicting outcome of military basic training for individuals referred for psychological evaluation. *Journal of Personality Assessment, 72,* 256-265. doi:10.1207 /S15327752JP720210

Celmer, V & Winer, J. L. (1990). Female aspirants to the Roman Catholic priesthood. *Journal of Counseling and Development, 69,* 178-183. Retrieved from http://www.counseling.org/Publications/Journals.aspx

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Psychology,*37*,* 319–338. doi:10.1016/S0092-6566(02)00578-0

Chamorro-Premuzic, T., & Furnham, A. (2006). Intellectual competence and the intelligent personality: A third way in differential psychology. *Review of General Psychology*, *10*(3), 251-267. doi: 10.1037/1089-2680.10.3.251

Chernyshenko, O.S. & Ones, D.S. (1999). How selective are psychology graduate programs? The effect of the selection ration on GRE score validity. *Educational and Psychological Measurement*, *59*(6), 951-961. doi:10.1177 /00131649921970279

Chibnall, J.T., & Detrick P. (2003). The Neo PI-R Inwald Personality Inventory and MMPI-2 In The Prediction Of Police Academy Performance: A Case for Incremental Validity. *American Journal of Criminal Justice, 27,* 233-248. doi:10.1007/BF02885696

Compton, K. (2009). *Using the MMPI-2 as a predictor of academic success* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses database. (UMI No. 3352025)

Costa, P.T., & McCrae, R.R. (1991). *The NEO Five Factor Inventory*. Odessa, FL: Psychological Assessment Resources.

Costa, P.T., & McCrae, R.R. (1992). *Revised NEO Personality Inventory and NEO Five Factor Inventory Professional manual.* Odessa; FL: Psychological Assessment Resources.

Costanzo, M. & Philpott, J. (1986). Predictors of therapeutic talent in aspiring clinicians: A multivariate analysis. *Psychotherapy, 23*, 363-369. doi:10.1037/h0085624

Daehnert, C., & Carter, J.D. (1987). The prediction of success in a clinical psychology graduate program. *Educational and Psychological Measurement*, *47*, 1113-1125. doi:10.1177/0013164487474029

Davis, D. & Widseth, J.C. (1978). A Minnesota Multiphasic Personality Inventory indicator of psychological distress in male students. *Journal of Counseling Psychology*, *25*(5), 469-472. doi:10.1037//0022-0167.25.5.469

Dao, T.K., Prevatt, F., Horne, H.L. (2008). Differentiating psychotic patients from nonpsychotic patients with the MMPI-2 and Rorschach. *Journal of Personality Assessment*, *90*(1), 93-101. doi:10.1080/00223890701693819

De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: A review. *European Journal of Personality, 10,* 303–336. doi:10.1002/(SICI)1099-0984(199612)10:5<303::AID-PER262>3.3.CO;2-U

Detrick, P., Chibnall, J.T., & Rosso, M. (2001). Minnesota Multiphasic Personality Inventory—2 in police officer selection: Normative data and relation to the Inwald Personality Inventory. *Professional Psychology: Research and Practice, 32*(5)*,* 484-490. doi:10.1037//0735-7028.32.5.484

Di Fabio, A. & Palazzeschi, L.(2009). An i-depth look at scholastic success: Fluid intelligence, personality traits or emotional intelligence? *Personality and Individual Differences*, *46*, 581-585. doi:10.1016/j.paid.2007.06.025

Dunnette, M.D., Bonas, D.A., & Bosshardt, M.J. (1981). *Electric Power Plant Study: Prediction of inappropriate, unreliable, or aberrant job behavior in nuclear power plant settings.* Minneapolis: Personnel Decisions Research Institute.

Edwards, W.R., & Schleicher, D.J. (2004). On selecting psychology graduate students: Validity evidence for a test of tacit knowledge. *Journal of Educational Psychology*, *96*(3), 592-602. doi:10.1037/0022-0663.96.3.592

Epstein, R.M. & Hundert, E.M. (2002) Defining and assessing professional competence *JAMA*.287:226-235. doi:10.1001/jama.287.2.226

Exner, J.E., Jr. (2003) *The Rorschach: A comprehensive system:Vol 1, Basic foundations* (4th ed.). New York :Wiley.

Farsides, T. & Woodfield, R. (2003). Individual differences and undergraduate academic success: The roles of personality, intelligence, and application. *Personality and Individual differences, 34*(7), 1225-1243. doi:10.1016/S0191-8869(02)00111-3

Forrest, L., Elman, N., Gizara, S., & Vacha-Haase, T. (1999). Trainee impairment: A review of identification, remediation, dismissal, and legal issues. *The Counseling Psychologist, 27,*627–686. doi:10.1177/0011000099275001

Frame, M. W., & Stevens-Smith, P. (1995). Out of harm's way: Enhancing monitoring of dismissal processes in counselor education programs. *Counselor Education and Supervision, 35,* 118–129. Retrieved from http://www.counseling.org/Publications/Journals.aspx

Furnham, A., & Monsen, J. (2009). Personality traits and intelligence predict academic school grades. *Learning and Individual Differences, 19,* 28-33. doi:10.1016/j.lindif.2008.02.001

Furnham, A., Monsen, J. & Ahmetoglu, G. (2009). Typical intellectual engagement, big five personality traits, approaches to learning and cognitive ability predictors of

academic performance. *British Journal of Educational Psychology*, *79*, 769-782. doi:10.1348/978185409X412147

Gall, M.D., Gall, J.P., & Borg, W.R. (2007). Educational Research, 8th Edition Pearson Education, Inc.

Garetz, F.K.; Anderson, R.W. (1973). Patterns of professional activities of psychiatrists: A follow-up of 100 psychiatric residents. *American Journal of Psychiatry*, *130*(9), 981-984. Retrieved from http://ajp.psychiatryonline.org

Gaubatz, M.D., & Vera, E.M. (2002). Do formalized gatekeeping procedures increase programs' follow-up with deficient trainees? *Counselor Education and Supervision, 41*(4), 294–305. Retrieved from http://www.counseling.org /Publications/Journals.aspx

Goldberg, E. & Alliger, G. (1992). Assessing the validity of the GRE for students in psychology: A validity generalization approach. *Educational and Psychological Measurement, 52*, 1019-1027. doi:10.1177/0013164492052004026

Graham, J.R. (1988). *Establishing validity of the revised form of the MMPI.* Symposium presentation at the 96th Annual Convention of the American Psychological Association, Atlanta, GA.

Graham, J.R. (2006) *MMPI-2 Assessing personality and psychopathology*. New York: Oxford University Press.

Green, S.B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research*, *26* (3), 499-510. doi: 10.1207 /s15327906mbr2603_7

Hair, J.F., Anderson, R.E., Tatham, R.L. & Black, W.C. (1995). Multivariate data analysis. Upper Saddle River, N.J.: Prentice-Hall.

Hargrave, G.E., Hiatt, D., Gaffney, T.W. (1988). F+4+9+Cn: An MMPI measure of aggression in law enforcement officers and applicants. *Journal of Police Science & Administration,16*, 268-273.

Haynes, S.N. & Lench, H.C. (2003). Incremental validity of new clinical assessment measures. *Psychological Assessment*, *15*, 4, 456-466. doi:10.1037/1040 -3590.15.4.456

Hiatt, D., & Hargrave, G. E. (1988). MMPI profiles of problem peace officers. *Journal of Personality Assessment. 52* (4), 722-731. Retrieved from http://0-www .informaworld.com.source.unco.edu/smpp/title~content=t775653663~db=all

Hirschberg, N., & Itkin, S. (1978). Graduate student success in psychology. *American Psychologist*, *33*, 1083-1093.

Holmes, C. (2003). One obvious reason doctoral programs in psychology use the graduate record examination. *Psychological Reports*, *92*, 692. doi:10.2466 /PR0.92.2.692

Holmes, C.B. & Beishline, M.J. (1996). Correct classification, false positives, and false negatives in predicting completion of the Ph.D. from GRE scores. *Psychological Reports*, *79*, 939-945. Retrieved from http://www.ammonsscientific.com

Hunsley, J. (2003). Introduction to the special section on incremental validity and utility in clinical assessment. *Psychological Assessment, 15*(4), 443-445. doi:10.1037 /1040-3590.15.4.443

Hunsley, J. & Meyer, G.J. (2003) The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment*, *15* (4), 446-455. doi: 10.1037/1040-3590.15.4.446

Jansen, D.G., Bonk, E.C., & Garvey, F.J. (1973). MMPI characteristics of clergymen in counseling training and their relationship to supervisors' and peers' ratings of counseling effectiveness. *Psychological Reports, 33,* 695-698. Retrieved from http://www.ammonsscientific.com

Johnson, W. & Campbell, B. (2004). Character and fitness requirements for professional psychologists: Training directors' perspectives. *Professional Psychology: Research & Practice*, 35, 405-411. doi:10.1037/0735-7028.35.4.405

Kelly, W.L. (1974). Psychological prediction of leadership in nursing. *Nursing Research, 23,* 83- 42. doi:10.1097/00006199-197401000-00009

Kerl, S.B., Garcia, J.L., McCullough, C. S., & Maxwell, M.E. (2002). Systematic evaluation of professional performance: Legally supported procedure and process. *Counselor Education & Supervision*, *41*, 321-334. Retrieved from http://www .counseling.org/Publications/Journals.aspx

King, A. R., & Bailly, M. D. (2002). MMPI-2 predictors of academic performance. *Journal of Psychoeducational Assessment, 20*(3), 258-267. doi:10.1177 /073428290202000303

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, *56* (5), 746-759. doi:10.1177 /0013164496056005002

Kirk, R. E. (2001). Promoting good statistical practices: Some suggestions. *Educational and Psychological Measurement, 61* (2), 213-218. doi:10.1177 /00131640121971185

Kodman, F. (1984). Some personality traits of superior university students. *Social Behavior and Personality, 12* (2), 135-138. doi:10.2224/sbp.1984.12.2.135

Kuncel, N.R., Crede, M. & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research*, *75*(1), 63-82. doi:10.3102 /00346543075001063

Kuncel, N.R. , Hezlett, S. A. (2010). Fact and fiction in cognitive ability testing for admissions and hiring decisions. *Current Directions in Psychological Science*, *19*(6), 339-345. doi:10.1177/0963721410389459

Kuyken, W., Peters, E., Power, M. J., & Lavender, T. (2003). Trainee clinical psychologists' adaptation and professional functioning: A longitudinal study. *Clinical Psychology &Psychotherapy, 10*, 41–54. doi:10.1002/cpp.350

Lachar, D., (1974). Prediction of early US Air Force freshman cadet adaptation with the MMPI. *Journal of Counseling Psychology, 21,* 404-408. doi:10.1037/h0037112

Lamb, D.H., Cochran, D.J., & Jackson, V.R. (1991). Training and organizational issues associated with identifying and responding to intern impairment. *Professional Psychology: Research and Practice*, *22*, 291-296. doi:10.1037//0735-7028 .22 .4.291

Leverett-Main, S. (2004). Program directors'perceptions of admission screening measures and indicators of student success. *Counselor Education & Supervision*, *43*, 207-219. Retrieved from ttp://www.counseling.org/Publications/Journals.aspx

Lievens, F., Dilchert, S. & Ones, D.S. (2009). Personality scale validities increase throughout medical school. *Journal of Applied Psychology*, *94* (6), 1514-1535. doi:10.1037/a0016137

Lounsbury, J. W., Sundstrom, E., Loveland, J. M., & Gibson, L. W. (2003). Intelligence, "Big Five" personality traits, and work drive as predictors of course grade. *Personality and Individual Differences, 35,* 1231–1239. doi:10.1016/S0191 -8869(02)00330-6

Lumadue, C. A., & Duffey, T. H. (1999). The role of graduate programs as gatekeepers: A model for evaluating student counselor competence. *Counselor Education and Supervision, 39,* 101–110 Retrieved from http://www.counseling.org /Publications /Journals.aspx.

McAdams, C. & Foster, V.A. (2007) A guide to just and fair remediation of counseling students with professional performance deficiencies. *Counselor Education & Supervision, 47*, 2-13. Retrieved from http://www.counseling.org /Publications /Journals.aspx

Miller,J.R. (2010) Lawsuit claims college ordered student to alter religious views on homosexuality or be dismissed. http://www.foxnews.com/us/2010/07/27/georgia -university-tells-student-lose-religion-lawsuit-claims/

Muller, B.P., & Bruno, L.N. (1988). *The effects of ethnicity on personality assessment in police candidates.* Paper presented at the 23rd Annual Symposium on Recent

Developments in the Use of the MMPI, St. Petersburg, FL.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J., Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*, 77-101. doi:10.1037//0003-066X.51.2.77

Noftle, E.E., & Robins, R.W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, *93* (1), 116-130. doi:10.1037/0022-3514.93.1.116

Norcross, J.C. & Kohout, J.L., & Wicherski, M. (2005). Graduate study in psychology 1971 to 2004. *American Psychologist*, *60*(9), 959 – 975. doi:10.1037/0003-066X .60.9.959

Olkin, R., & Gaughen, S. (1991). Evaluation and dismissal of students in master's level clinical programs: Legal parameters and survey results. *Counselor Education and Supervision, 30,* 276–288. Retrieved from http://www.counseling.org/Publications /Journals.aspx

Osborne, J.W. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical Assessment, Research & Evaluation, 15*(12), 1-7. Retrieved from http://pareonline.net/

Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology, 89* (2), 187–207. doi: 10.1037/0021-9010.89.2.187

Ozer, D.J., & Benet-Martinez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review Psychology*, *57*, 401-421.doi:10.1146/annurev.psych.57.102904.190127

Pallant, J. (2007). SPSS Survival Manual. Open University Press : New York, NY.

Petrocelli, J.V. (2003). Hierarchical multiple regression in counseling research: Common problems and possible remedies. *Measurement and Evaluation in Counseling and Development*, *36*, 9-22. Retrieved from http://0-ec.sagepub.com.source.unco.edu/

Plaintiff v. Rector and Board of Visitors of The College of William and Mary, No. 42–1983 (E. Dist. August 26, 2003).

Poropot, A.E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135* (2), 322-338. doi:10.1037/a0014996

Rouse, S. V. & Butcher, J. N. (1995). *Annotated bibliography on the use of the MMPI/MMPI-2 in personnel and educational selection.* Minneapolis, MN: University of Minnesota Press.

Shusman, E.J., Inwald, R.E., & Kantz, H.F. (1987). A cross validation study of police recruit performance as predicted by the IPI and MMPI. *Journal of Police Science and Administration, 15*, 162-169.

Smaby, M.H., Maddux, C.D., Richmond, A.S., Lepkowski, W.J. & Packman, J. (2005). Academic admission requirements as predictors of counseling knowledge, personal development and counseling skills. *Counselor Education & Supervision*, *45*, 43-57. Retrieved from http://www.counseling.org/Publications/Journals.aspx

Smith, H. L., Robinson, E. H., & Young, M. E. (2007). The relationship among wellness, psychological distress, and social desirability of entering master's-level counselor

trainees. *Counselor Education and Supervision, 47,* 96-109. Retrieved from http://www.counseling.org/Publications/Journals.aspx

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence.* New York: Cambridge Univ. Press.

Sternberg, R.J. & Williams, W.M. (1997). Does the graduate record examination predict meaningful success in the graduate training of psychology? A case study. *American Psychologist, 52*, 630-641. doi:10.1037//0003-066X.52.6.630

Strupp, H.H.& Bloxom, A.L. (1975). An approach to defining a patient population in pscyhotherapy research. *Journal of Counseling Psychology*, *3*, 231-237. doi:10.1037/h0076881

Trapmann, S., Hell, B., Hirn, J.W., & Schuler, H. (2007). Meta-Analysis of the relationship between the big five and academic success at university. *Journal of Psychology*, *215*, 132-151. doi.10.1027/0044-3409.215.2.132

Vacha-Haase, T., Davenport, D.S., & Kerewsky, S.D. (2004). Problematic students: Gatekeeping practices of academic professional psychology problems. *Professional Psychology: Research and Practice, 35*, 115–122. doi.10.1037 /0735-7028.35.2.115

Weiner, I.B., & Greene, R.L. (2008). Handbook of personality assessment. New York: Wiley.

Weiss, W.U., Serafino, G., Serafino, A., Willson, W. & Knoll, S. (1998). Use of the MMPI-2 to predict the employment continuation and performance ratings of recently hired police officers. *Journal of Police and Criminal Psychology*, *13*(1), 40-44. doi:10.1007/BF02819065

Weiss, W.U., Serafino, G., Serafino, A., Willson, W., Sarsany, J. & Felton, J. (1999). Use of the MMPI-2 and the Inwald Personality Inventory to identify the personality characteristics of dropouts from a state police academy. *Journal of Police and Criminal Psychology*, *14* (1), 38-42. doi:10.1007/BF02813861

Wetzler, S., Khadivi, A., & Moser, R. (1998). The use of the MMPI-2 for the assessment of depressive and psychotic disorders. *Assessment*, *5*(3) 249-261. doi:10.1177 /107319119800500305

Wilkerson, K. (2006). Impaired students: applying the therapeutic process model to graduate training programs. *Counselor Education & Supervision*, *45*, 207-217. Retrieved from http://www.counseling.org/Publications/Journals.aspx

Zeidner, M., & Matthews, G. (2000). Intelligence and personality. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 581–610). New York: Cambridge University Press.

APPENDIX A

JOURNAL MANUSCRIPT

Abstract

The prediction of success in graduate education is of growing interest in the field of professional psychology. Notably there is debate as to the most effective variables to build predictive models and how to define success. This study analyzed whether the prediction of performance in a professional counseling masters program (PCMP) using undergraduate grade point average and Pre-Admissions Workshop rating scores could be incremented by adding personality traits, as measured by the MMPI-2. Graduate grade point average (GGPA) and internship evaluation ratings (IER) were both used as success criteria in order to investigate whether a stronger predictive model could be built using a traditional outcome variable or a criterion more in line with the training objective of a professional program. The two hierarchical regression analyses produced a number of key findings. Across both analyses, the MMPI-2 independently accounted for the most variance in performance outcome, after controlling for the traditional admissions variables. The final model accounted for 24.4% ($p<.05$) of the variance in GGPA and of that; the MMPI-2 uniquely explained 17.7% of the variance. When IER was used as the criterion, the final model accounted for 15.6% ($p<.10$) of the variance, and 14.5% of that variance was uniquely explained by the MMPI-2. Overall, these findings provided empirical support for the inclusion of an objective, valid personality instrument in PCMP admissions procedures and suggest that personality characteristics are veritably relevant to academic performance.

Keywords: Personality, MMPI-2, incremental validity, academic admission criteria, academic performance prediction, professional counseling masters programs

Introduction

Over the past twenty years, there has been considerable focus on competency in mental health literature. This is especially evident in relation to screening trainees for admissions and assessing individuals once they have been accepted into professional psychology graduate programs. The definition of competence utilized by the American Psychological Association's Assessment of Competency Benchmarks Work Group (2007) is as follows, "professional competence is the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values and reflection in daily practice for the benefit of the individual and community being served"

(Epstein & Hundert, 2002, p.226). To be a competent psychology graduate student, therefore, requires certain coalescence in the dimensions of ability, aptitude and application. As is the case with all professions, the definition and measurement of those constructs is both philosophically complex and debatable.

Currently, there is a lack of consensus as to the best predictors of success in professional counseling masters programs (PCMP) and debate about how to define "success." In order to assess whether an applicant has the potential to succeed in a PCMP, admissions committees commonly utilize both objective and subjective methods to screen applicants. Objective methods tend to be quantitative, cognitive measures such as the Graduate Record Examination (GRE) and undergraduate grade point average. Conversely, personal characteristics tend to be assessed in a more qualitative, subjective fashion through impressions from letters of recommendation, personal interviews or essays.

In a national survey amongst clinical training directors, Johnson and Campbell (2004) found that clinical training directors are concerned about the moral character and psychological fitness of applicants and that there is dissatisfaction related to the lack of valid and reliable measures for pre-admissions assessment. Constanzo and Philpott (1986) found that interpersonal factors predicted therapeutic talent the most and maintain that traditional admissions criteria reflect an outdated view of intelligence because it provides a limited scope of intelligence heavily weighted on purely cognitive abilities. In fact, there seems to be a consensus in the literature that there are multiple intelligences that contribute to success as a professional counselor and that it is a shortcoming of admissions criteria that these be overlooked (Sternberg & Williams, 1997). Wheeler

(2000) conducted an investigation about the attributes of a "good counselor" and found that personality factors were the most significant discriminate factors differentiating "good verse bad" counselors. Therefore, the purpose of the present study was to determine if adding an objective personality assessment could strengthen the prediction of success beyond what could be predicted with more traditional admissions criteria. Additionally, the present study attempted to determine if the predictive strength of the model can be improved with the use of a criterion more in line with the specific training objectives of a PCMP.

MMPI-2 and Performance Research

Although the MMPI was originally intended for use with clinical populations, one of its more current applications has been to screen for sensitive occupations and for students in particular training programs. These occupations are associated with psychological factors that would significantly impact performance, including those that are high stress, involve personal risk and require personal responsibility (Butcher, 1985, 1991). In terms of appropriateness of use in this regard, Butcher, Ones and Cullen (2006) state that more than 570 books and articles have empirically supported the use of the MMPI/MMPI-2 in personnel and educational applications (Rouse & Butcher, 1995, as cited in Butcher, 2006, pp. 382-383). Literature on the ability of the MMPI/MMPI-2 to predict actual job performance, however, is more limited and complicated to determine.

There are several obstacles that make it difficult to determine the empirical validity of the MMPI-2 as a predictor of job and graduate school performance. First, if the MMPI-2 was used as part of the prescreening hiring process, it is likely that applicants with profiles suggestive of serious pathology would not be hired. This significantly limits the ability to study deviant profiles and actual job performance, and

restricts the range of scores (i.e., only scores in the normative range would be available). Similarly, if the MMPI-2 accurately identified applicants who would not be suitable for sensitive occupations and training programs, it would follow that there would be fewer individuals who would make it through the screening process and could possibly be identified later as exhibitors of poor job performance. Consequently, it can be challenging to obtain a large enough sample size of poor performers to make meaningful statistical interpretations (Hiatt & Hargrave, 1988). Therefore, it could be argued that the better the MMPI-2 performs at personnel screening the harder it could be to empirically validate the results.

A review of the literature demonstrated that the MMPI and the MMPI-2 have been used extensively in the selection of police officers (Butcher, 1979; Detrick, Chibnall & Rosso, 2001;Weiss, Serafino, Serafino, Willson & Knoll, 1998; Weiss, Serafino, Serafino, Willson, Sarsany & Felton 1999) and numerous studies have been conducted on the validity of the MMPI-2 in predicting their future job performance (Chibnall & Detrick, 2003; Hargrave & Hiatt, 1987; Hargrave, Hiatt & Gaffney, 1988; Shusman, Inwald & Knatz, 1987). A review of other studies conducted to confirm the efficacy of the MMPI and MMPI-2 for personnel screening of sensitive occupations include the following: nurses (Kelly, 1974), nuclear power plant employees (Dunnette, Bownas & Bosshardt, 1981), military personnel (Callan, 1997; Carbone, Cigrang, Todd and Fielder, 1999), clergy (Banks, Mooney, Mucowski & Williams, 1984; Celmer & Winer, 1990; Jansen, Bonk & Garvey, 1973), and psychiatric residents (Garetz & Anderson, 1973).

In regard to the prediction of academic performance, Butcher and Rouse (1996) maintained that the MMPI-2 does detect individuals who are likely to experience

emotional maladjustment in a stressful academic environment. Although recent research to support that assertion is sparse, there is evidence that personality characteristics as measured by the MMPI and MMPI-2 can predict academic performance. Lachar (1974) studied a freshman class at the United States Air Force Academy (N=1,389) and found the rates of attrition and emotional maladjustment were significantly higher for the group of cadets identified as being "high risk" when screened with the MMPI. Strupp and Bloxom (1975) compared male undergraduate college students with T scores above 60 on Scale 2 (Depression) and Scale 7 (Psychasthenia) with a random group of undergraduate students over the course of eight-and-a-half years. Results indicated that the students with high scores on Scale 2 and Scale 7 had more emotional and behavioral problems, as well as longstanding, negative concerns related to academic, personal and occupational performance. A similar study by Davis and Widseth (1978) supported the findings by Strupp and Bloxom (1975). Other notable studies examining the associations between the MMPI/MMPI and academic performance include: Kodman (1984), King and Bailly (2002) and Compton (2009).

Predictive Validity, Incremental Validity and Criterion Issues

There is a lack of consensus in the literature as to what are the most valid predictive measures to utilize for admissions criteria to graduate psychology programs, as well as debate regarding a definition for the criterion of "success." More recently, research on academic performance, in general, has focused on incremental validity rather than approaching prediction from the standpoint of one independent variable being superior to another. "Incremental validity is defined as the degree to which a measure explains or predicts a phenomenon of interest, relative to other measures" (Haynes &

Lench, 2003, p. 456). In applied psychology, improved prediction can mean several things, such as increased specificity or the power and ability to make decisions. The emphasis is on the extent to which a variable adds to the prediction model beyond what could have been predicted with other variables (Hunsley, 2003), or simply utilizing base rate data (Hunsley & Meyer, 2003).

In addition to distinguishing that new data can strengthen an existing statistical model, incremental validity research can also determine how well a variable accounts for outcome variance in comparison to the other measures (Haynes & Lench, 2003). It can provide evidence of whether multiple sources of data should be jointly considered, irrespective of whether individually the variable is known to contribute to the prediction of a criterion. Hunsley (2003) stated, "Any scientific justification for obtaining and combining multiple sources of data can only come from research evidence in which unique and shared variance among measures is directly examined with respect to a criterion of interest" (p. 443). Therefore, incremental validity research can assist in the development of a more parsimonious model because it can determine if the addition of multiple predictor variables improves the statistical model beyond what variance could have been explained by fewer variables. Utility, finally, is also an important consideration that can be addressed through incremental validity research because it can be used as evidence to support or refute the use of additional variables based on elements such as time and cost (Hunsley, 2003).

Predictive Validity: Intelligence, Personality and Academic Performance

There is consensus in the literature that intelligence is a strong predictor of academic success and that the predictive validity of future academic performance

decreases as individuals progress to higher levels of education (Busato, Prins, Elshout & Hamaker, 2000; Di Fabio & Palazzeschi, 2009; Farsides & Woodfield, 2003; Furnham & Monsen, 2009; Trapmann, Hell, Hirn & Schuler, 2007). This is especially the case with graduate school; however, there is a lack of predictive validity research at this level of education. Large scale studies as well as meta-analyses have evidenced that the prediction of academic performance using cognitive variables (e.g., GPA or academic achievement tests) can be incremented by adding objective measures of personality (Lievens et al., 2009; Poropat, 2009; Trapmann et al., 2007). In fact, Poropat (2009), who conducted the most comprehensive meta-analytic investigation of personality-academic performance relationships concluded, "Personality should take a more prominent place in future theories of academic performance and not merely as an adjunct to intelligence" (p.333).

It is important to briefly mention that over the past twenty years, empirical research on the relationship between personality and academic performance has burgeoned, and a review of the literature demonstrated that researchers have predominantly used the Big Five Factor Model (FFM) of personality as the taxonomy for classifying personality dimensions. For example, personality traits as measured with instruments based on the Five Factor Model have been found to add incremental validity in predicting a range of academic outcomes, such as final grades and grade-point average at various levels of education (Chamorro-Premuzic & Furnham, 2003; Lievens, Dilchert &Ones, 2009; Poropot, 2009). That being said, it would be misleading to assume that the volume of research using the FFM, and not other valid, reliable and objective personality instruments, reflects a consensus in the field of professional psychology that the FFM is

the best method for understanding and measuring dimensions of personality. Chamorro-Premuzic and Furnham (2006) addressed this trend and explained, "The snowball effect regarding the consensus on the Big Five is indicative of a self-fulfilling prophecy: The more researchers agree on the validity of the Big Five, the more research is conducted to validate the Big Five" (p. 256). Therefore, it is important to underscore that the reason research using the FFM has dominated the literature is because there is a significant lack of recent research that has used other personality assessments, such as the MMPI-2.

Thus, the purpose of this study was twofold. Given that there is a lack of consensus as to whether personality characteristics should be objectively assessed as part of standard admissions to graduate school, the first objective of the present study was to provide empirical evidence to support the inclusion of valid, reliable personality instruments in professional counseling masters programs (PCMP) admissions procedures. This was accomplished by testing the hypothesis that the MMPI-2 can predict final grade point average and provide incremental validity above what can be explained by previous academic achievement; namely undergraduate grade point average. The multiple regression analysis included ratings given by the admissions committee members during the Pre-Admissions Workshop in order to test whether adding an objective personality instrument could strengthen the predictive validity of standard admissions criteria, or if a qualitative rating was a stronger predictor of future success in the PCMP.

The second goal of the present study was to meet the need cited in the literature to test criteria other than GPA as a measure of success and overcome the limitations of using GPA as the criterion in a graduate program where application of knowledge becomes a quintessential indicator of success. This was especially significant for this

sample where the variance in performance on practicum and internship was not reflected in GPA scores because it was factored in dichotomously as either satisfactory or unsatisfactory. For example, one might have to repeat a practicum or struggle in an area of internship (e.g., openness to feedback) but that individual's struggle to demonstrate competence was not be reflected numerically in GPA. That variance in performance was captured on internship evaluations, which was one of the reasons a separate regression analysis was conducted using the same predictor variables and scores on internship as the criterion.

Additionally, internship evaluation scores were analyzed because the form is standardized and the literature suggested that it was a more accurate reflection of performance in a PCMP (Daehnert and Carter, 1987). Therefore, the second analysis contributed to the literature and benefited PCMPs in a similar manner as the first. The main purpose; however, was to test the amount of variance that can be explained in performance when using a criterion more aligned with the training objectives of a PCMP.

Research Hypotheses

H1 The prediction of graduate grade point average in a professional counseling masters program using undergraduate grade point average and Pre-Admissions Workshop rating scores can be incremented by adding personality traits, as measured by the MMPI-2.

H2 The prediction of internship evaluation ratings in a professional counseling masters program using undergraduate grade point average and Pre-Admissions Workshop rating scores can be incremented by adding personality traits, as measured by the MMPI-2.

H3 Personality traits, as measured by the MMPI-2, will have more predictive validity when internship evaluations rather than graduate grade point averages are used as the criterion.

Method

To test H1 and H2, two separate hierarchical multiple regression analyses were conducted. The entry of the predictor variables were done in blocks, respectively as follows: undergraduate GPA, mean score from Pre-Admissions Workshop ratings and *T*-scores on three validity scales and ten clinical scales of the MMPI-2. The criterion variable was graduate GPA in the first analysis and the mean score from the internship evaluation in the second analysis. The third hypothesis was tested by comparing the total percentage of variance explained by personality traits in each predictive model.

Participants

The total sample included 142 individuals who graduated from a Professional Counseling Masters Program (PCMP) at a medium, public university in the Rocky Mountain region. The participants represented six consecutive graduating classes who graduated between 2003 and 2009. The participants had been selected for admissions into the program based on the following criteria: personal statement, GRE or GPA scores, three letters of recommendation, MMPI-2 assessment data, and ratings from a Pre-Admissions Workshop.

Participants, who met the requisite criteria, were admitted to the PCMP and completed an internship at the end of their training. The minimum requirements for all internships included the following: minimum of 300 direct client contact hours; minimum 600 on-site hours and minimum of one hour per week of individual supervision throughout duration of internship to be conducted by the participant's on-site supervisor. All participants who graduated from the program obtained a standardized evaluation from

their internship-site supervisor which was completed and returned to the University clinical director at the end of the participant's internship.

Procedures

The data were archival data spanning from years 2003-2009. In order to protect the confidentiality and anonymity of participants, official university staff extracted all data from participant's files and de-identified the data prior to being received by the primary investigator. The following data were extracted from the participants' respective academic files: demographic information, undergraduate GPA, mean rating score from Pre-Admissions Workshop, MMPI-2 scores on the 13 validity and clinical scales, rating scores on internship evaluations and final GPA.

Variables

All predictor variables were components of data assessed as part of the pre-admissions process at this PCMP. Given that applicants were not required to take the Graduate Record Examination (GRE) unless their undergraduate grade point average (UGPA) was below 3.0, UGPA was used as the cognitive predictor variable. Students were also required to take the MMPI-2 and participate in a Pre-Admissions Workshop. These scores were used as the non cognitive predictor variables. Specifically, three validity scale scores and ten clinical scale scores on the MMPI-2 as well as the mean score from faculty ratings at the Pre-Admissions Workshop. The rationale for the selection of these variables was to analyze whether the prediction of academic performance could be incremented by adding non cognitive predictor variables. The rationale for using two non-cognitive predictor variables was that it allowed for the examination of whether an objective, quantitative method of screening applicants or a

subjective, qualitative assessment made by the admissions committee had more incremental validity. Overall, analyzing these variables determined the utility of current admission procedures at a PCMP.

The criterion variable in the first analysis was graduate grade point average (GGPA). It was selected because it is the most commonly used criterion in predictive validity research in academia (Kuncel, Crede, & Thomas, 2005) which could aid in the generalizability of the findings. Internship evaluation mean scores were used as the criterion in the second analysis in order to test whether a criterion more in line with the PCMP training objectives could strengthen the predictor-criterion relationship in the statistical model. Additionally, using internship evaluations incorporated previous research. For example, Daehnert and Carter (1987) concluded that internship evaluations were the "ideal criteria" for measuring success in a PCMP and research by Leverett-Main (2004) indicated that program directors in CACREP accredited PCMP endorsed practicum/internship evaluations as the best measure of success in the program. The internship evaluations measured seven domains of professional competency on a five point Likert scale. The overall mean score were computed and then analyzed as the data from those evaluations.

Instruments

The Minnesota Multiphasic Personality Inventory-2 (MMPI-2). The MMPI-2 is the most popular assessment for psychopathology and clinical research and is the most widely used personality test throughout the world (Butcher, 2006; Butcher & Rouse, 1996; Graham, 2000). It is a self report inventory and a broadband measure of personality characteristics (Weiner & Greene, 2008) consisting of 567 true or false statements. It

takes between 60-90 minutes to complete, requires a 6th to 8th grade reading level and is appropriate for individuals eighteen and older. In the basic profiles, the MMPI-2 contains nine validity scales and ten clinical scales.

Internship Student Evaluation. The faculty of the professional counseling masters program adopted a standardized evaluation form to be scored at the completion of the graduate student's 600 hour internship. The Internship Student Evaluation assesses the following areas of counselor competency: Opening and Rapport, Interaction and Interview Skills, Counselor Responses, Counseling Relationship, Client Conceptualization, Termination, and Case Conceptualization and Supervision. There are a total of 22 ratings on the evaluation which are scored on a 5 point Likert-type scale. There is also the option of scoring N/A if the behavior was not observed by the on-site supervisor (see Appendix C).

Results

Descriptive Statistics

The gender of the participants was 15.5% male and 84.5% female. At the time of graduation, participants' ranged in age from 25 to 60 years-old and the mean age was 36 years-old. Only 4.6% of the participants did not report their age in the admission's packet; however, 25.4% did not report their ethnicity. Of those who did endorse ethnicity, 62% identified as White/Caucasian, 2.8% identified as Black/African American, 7.0% identified as Hispanic/Latino and 2.8% identified as Asian/Pacific Islander.

The means and standard deviations for Graduate Grade-Point Average (GGPA), Internship Evaluation Scores (IES), Undergraduate Grade-Point Average (UGPA), Pre-Admissions Workshop Rating score (PWR) and MMPI-2 scores are reported in Table 1.

It is notable that all mean T scores on the MMPI-2 of the participants were within one standard deviation of the normative sample mean scores of male medical/psychology applicants.

Table 1

Descriptive statistics for predictor and outcome variables

Variable	*M*	*SD*	Minimum	Maximum
GGPA	3.87	.17	3.27	4.0
IES	4.53	.46	2.57	5.0
UGPA	3.45	.39	2.10	4.0
PWR	3.44	.54	1.00	4.0
MMPI-2 *T* Scores				
L	53.91	8.87	38	76
F	43.82	6.05	36	82
K	61.23	7.37	41	76
Hs (1)	48.96	6.15	38	69
D (2)	44.08	5.56	30	64
Hy (3)	51.87	6.92	38	69
Pd (4)	53.61	8.51	37	84
Mf (5)	53.49	9.97	30	74
Pa (6)	50.92	6.98	34	70
Pt (7)	47.92	5.79	32	64
Sc (8)	48.79	6.41	31	73
Ma (9)	49.96	8.29	35	79
Si (10)	40.53	6.46	30	61

Bivariate Pearson correlation coefficients for all variables in the analysis testing H1 are presented in Table 2. As shown, GGPA scores were significantly correlated with UGPA and Scales 2, 5, 8 and 9. Table 3 shows bivariate Pearson correlation coefficients for all variables in the analysis testing hypothesis two. As shown, only Scale 2 significantly correlated with internship evaluation ratings. These correlation matrices were also examined for multicollinearity. Given that there were no two independent variables with a bivariate correlation of .7 or higher, all independent variables were retained in the model (Pallant, 2007).

Table 2

Correlations Among Graduate Grade Point Average, Undergraduate Grade Point Average, Pre-Admissions Workshop Rating and MMPI-2 Variables (Hypothesis 1)

Variable	GGPA	UGPA	PWR	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si
GGPA																
UGPA	.26*															
PWR	-.06	-.09														
L	-.52	-.12	-.13													
F	-.12	.01	.19*	-.19*												
K	.07	.09	-.02	.44*	-.15											
Hs (1)	-.11	.03	.17*	.10	.25*	.39*										
D (2)	-.19*	-.04	.14	-.06	.34*	-.08	.49*									
Hy (3)	.005	.08	.12	.12	.11	.26*	.69*	.35*								
Pd (4)	-.02	.10	.11	-.06	.44*	.21*	.39*	.35*	.40*							
Mf (5)	-.21*	-.07	-.14	.13	.00	.19*	.13	.05	-.03	.02						
Pa (6)	.05	.12	.13	-.08	.23*	-.06	.19	.16	.13	.24*	-.13					
Pt (7)	-.01	.08	.03	-.05	.34*	.31*	.43	.40*	.36*	.48*	.10	.29				
Sc (8)	-.17*	.07	.15	-.00	.50*	.39*	.57	.29*	.38*	.60*	.17	.26	.64			
Ma (9)	-.28*	-.70	-.07	-.18*	.29*	-.32*	-.04	-.02	-.04	.20*	.01	.13	.13	.31*		
Si (10)	-.04	-.05	.10	-.05	.25*	-.30*	.09	.30*	-.15	-.02	-.04	-.00	.14	-.03	-.22*	

*indicates significance of .05 or less

Table 3

Correlations Among Internship Evaluation Rating, Undergraduate Grade Point Average, Pre-Admissions Workshop Rating and MMPI-2 Variables (Hypothesis 2)

Variable	IER	UGPA	PWR	L	F	K	Hs	D	Hy	Pd	Mf	Pa	Pt	Sc	Ma	Si
IER																
UGPA	.10															
PWR	.05	-.09														
L	.07	-.12	-.13													
F	.09	.01	.19*	-.19*												
K	.07	.09	-.02	.44*	-.15											
Hs 1)	.02	.03	.17*	.10	.25*	.39*										
D (2)	-.21*	-.04	.14	-.06	.34*	-.08	.48*									
Hy(3)	.01	.08	.12	.12	.11	.26*	.69*	.35*								
Pd (4)	.02	.10	.11	-.06	.44*	.21*	.39*	.35*	.40*							
Mf (5)	-.01	-.07	-.14	.13	.00	.19*	.13	.05	-.03	.02						
Pa (6)	.12	.12	-.13	-.08	.23*	-.06	.19*	.16	.13	.24*	-.13					
Pt (7)	-.08	.08	.03	-.05	.34*	.31*	.43*	.40*	.36*	.48	.10	.29				
Sc (8)	-.03	.07	.15	-.002	.50*	.39*	.57*	.29*	.38*	.60	.17*	.26*	.64*			
Ma(9)	-.12	-.07	-.08	-.18*	.29*	-.32*	-.04	-.02	-.04	.20*	.01	.13	.13	.31*		
Si (10)	-.06	-.05	.10	-.05	.25*	-.30*	.09	.30*	-.15	-.02	-.04	-.003	.14	-.03	-.22*	

*indicates significance of .05 or less

Hypothesis 1

Hierarchical multiple regression analysis was used to determine the incremental validity of the MMPI-2 in the prediction of GGPA in a PCMP after controlling for more commonly used criteria, undergraduate GPA (UGPA) and Pre-Admissions Workshop rating (PWR) scores. UGPA was entered at the first block, explaining 6.6% of the variance in GGPA, $F(1,140) = 9.97$, $p = .002$ ($R^2 = .066$, $\Delta R^2 = .066$). After entering PWR

scores in block two, the total variance explained by the model as a whole was 6.7%, $F(2,139)= 4.99$, $p=.008$ ($R^2 = .067$, $\Delta R^2 = .001$). When the MMPI-2 scores were added in step three, the final model accounted for 24.4% of the variance, $F(15,126) = 2.80$, $p=.001$ ($R^2 = .244$, $\Delta R^2= .177$). See Table 4 for the hierarchical regression analysis results. In the final model, three variables were statistically significant. UGPA had the lowest beta value (*beta*=.188, $p=.021$), then Scale 2 on the MMPI-2 (*beta* = -.237, $p=.025$) and then Scale 9 (*beta* = -.306, $p=.009$). Given that the MMPI-2 contributed an additional 17.7% of the variance above and beyond what was predicted by UGPA and Pre-Admissions Workshop ratings, H1 was supported.

Table 4

Hierarchical regression analysis of GGPA on predictor variables (Hypothesis 1)

Variable	*B*	*SE*	*Beta*	*t*	*p* value
Step 1 (R^2 = .066, ΔR^2= .066)					
Constant	3.49	.121		28.76	.000
UGPA	.110	.035	.26	3.16	.002
Step 2 (R^2 = .067, ΔR^2= .001)					
Constant	3.47	.14		24.14	.000
UGPA	.11	.04	.26	3.11	.002
PWR	.007	.03	.02	.29	.770
Step 3 (R^2 = .244, ΔR^2= .177)					
Constant	4.28	.35		12.24	.000
UGPA	.08	.03	.19	2.34	.021*
PWR	.006	.03	.02	.23	.820
L	-.001	.002	-.06	-.59	.557
F	.000	.003	.02	.16	.876
K	-.001	.003	.03	-.21	.838
1	-.001	.004	-.02	-..13	.894
2	-.007	.003	-.24	-2.26	.025*
3	.001	.003	.03	.19	.848
4	.002	.002	.10	.92	.358
5	-.003	.001	-.15	-1.81	.072
6	.001	.002	.05	.53	.598
7	.005	.003	.19	1.55	.124
8	-.005	.004	-.18	-1.18	.241
9	-.006	.002	-.31	-2.66	.009*
10	-.002	.003	-.07	-.64	.523

Note. N = 142.
*indicates significance of .05 or less

Hypothesis 2

Hierarchical multiple regression analysis was also used to determine the incremental validity of the MMPI-2 in the prediction of performance on internship evaluations in a professional counseling masters program after controlling for more commonly used criteria, undergraduate GPA and Pre-Admissions Workshop rating scores. Undergraduate GPA was entered at the first step, explaining .9% of the variance in internship evaluation ratings, F (1,140) = 1.28, p =.260 (R^2 = .009, ΔR^2 = .009). After

entering Pre- Admissions Workshop rating scores in block two, the total variance explained by the model, as a whole, was 1.1%, $F(2,139) = .764$, $p = .468$ ($R^2 = .011$, $\Delta R^2 = .002$). When the MMPI-2 scores were added in step three, the final model accounted for 15.6% of the variance, $F(15,126) = 1.554$, $p = .096$ ($R^2 = .156$, $\Delta R^2 = .145$). Although the additional incremental validity of the MMPI-2 was substantial, H2 was not significant at the .05 alpha level set for the study ($p = .096$). The model was significant at the .10 level and those results are presented in Table 5. It is also notable that the *p*-values for the Type III Sum of Squares were significant at the .05 level for MMPI-2 Scale F and Scale 2. This finding indicated that the unique contribution of each of those independent variables was significant after adjusting for the variability explained by all other independent variables.

Table 5

Hierarchical regression analysis of Internship Evaluation Score on predictor variables (Hypothesis 2)

Variable	*B*	*SE*	*Beta*	*t*	*p* value
Step 1 (R^2 = .009, ΔR^2 =.009)					
Constant	4.144	.347		11.955	.000
UGPA	.113	.100	.095	1.131	.260
Step 2 (R^2 = .011, ΔR^2 = .002)					
Constant	4.034	.410		9.836	.000
UGPA	.108	.100	.091	1.076	.284
PWR	.037	.073	.043	.505	.614
Step 3 (R^2 = .156, ΔR^2 = .145)					
Constant	4.720	1.026		4.603	.000
UGPA	.075	.101	.063	.740	.461
PWR	.020	.075	.023	.265	.792
L	.005	.005	.089	.874	.384
F	.019	.008	.254	2.306	.023*
K	-.004	.010	-.068	-.417	.677
1	.016	.012	.208	1.342	.182
2	-.029	.009	-.351	-3.163	.002*
3	-.003	.009	-.041	-.307	.759
4	.004	.006	.077	.681	.497
5	.002	.004	.033	.376	.707
6	.008	.006	.127	1.377	.171
7	-.002	.010	-.026	-.206	.837
8	-.011	.011	-.151	-.960	.339
9	-.010	.007	-.186	-1.527	.129
10	-.007	.008	-.096	-.809	.420

Note. N = 142.
*indicates significance of .05 or less

Hypothesis 3

Originally, this hypothesis was to be tested by comparing the total percentage of variance explained by personality traits in each statistical model. Given that the MMPI-2 explained 17.7% of the unique variance in GGPA and 14.5% of the unique variance in the performance on internship evaluation ratings; Hypothesis 3 was not supported. However, the main reason this hypothesis was rejected was because contrary to the

prediction of GGPA, none of the models (overall p values) were statistically significant predictors of performance on internship evaluations.

Discussion

The purpose of this study was to investigate whether there is empirical support to include an objective measure of personality characteristics as part of standard admissions procedures in a professional counseling masters program (PCMP). As first hypothesized, the MMPI-2 provided incremental validity in the prediction of graduate grade point average (GGPA) above what could be predicted by undergraduate grade point average (UGPA) and Pre-Admissions Workshop ratings (PWR). In fact, results from the hierarchical multiple regression analysis indicated that the overall model accounted for 25% of the variance in GGPA and of that, the MMPI-2 independently explained 17% of the variance. UGPA was significantly associated with GGPA ($r = .26$; $p < .05$) and produced a statistically significant change in R^2 (*beta* .188, $p = .021$; $p < .05$) in the analysis. Consistent with the Pearson correlation coefficients, PWR was not significantly associated with GGPA and did not significantly contribute to the model.

In the first model, it was estimated that a one standard deviation increase in UGPA would yield a .19 increase in GGPA. Given the consensus in the literature that cognitive measures can predict academic outcomes (Burton & Ramist, 2001; Furnham, Monsen & Ahmetoglu, 2009; Trapmann et al 2007) this study lends further support. Furthermore, it upholds the notion that when building a model to predict academic outcomes it is prudent to include both cognitive and non cognitive variables. For example, empirical support from meta-analyses and large scale studies have evidenced that the prediction of academic performance using cognitive variables (e.g., GPA or

academic achievement tests) can be incremented by adding objective measures of personality (Lievens et al., 2009; Poropat, 2009; Trapmann et al., 2007).

Conversely, the non cognitive variable, PWR, did not significantly contribute to the prediction of GGPA (ΔR^2 =.001). This finding brings into question the utility of subjective personality assessment as part of admissions screening, given research indicated that subjective methods are the most common admissions method of screening for personality or character currently used by PCMPs. However, it is also important to underscore that, in general, it may not be that subjective assessments of personality characteristics are poor predictors of performance in a PCMP; rather it could have been this particular variable performed poorly. In other words, it is likely that since PWRs were based on a 4 point Likert-type rating scale and the majority of participants (i.e., admitted students) scored between three and four on the rating scale; restriction of range limited the predictive ability of this variable in this analysis.

Although Pearson correlation coefficients indicated that Scales 2 (D), 5 (Mf), 8 (Sc) and 9 (Ma) significantly correlated with GGPA ($p<.05$), regression analysis results showed Scale 9 (Hypomania) and 2 (Depression) were the significant contributors to the prediction of GGPA. In fact, a one standard deviation decrease in score on the Hypomania scale would yield a .306 increase in GGPA, suggesting that individuals with higher scores on the Hypomania scale tended to have lower GGPA scores. This finding was consistent with extant literature on predictor – criterion relationships using the MMPI/MMPI-2 and performance outcomes (Compton, 2009; Hiatt & Hargrave, 1988; King & Bailey, 2002). For example, the Ma scale predicted undergraduate grade point average in Compton's (2009) study, it was shown to predict "problem" officers by Hiatt

and Hargrave (1988) and King and Bailey (2002) found that high scores correlated with poor academic performance as measured by undergraduate GPA.

As would be predicted from similar studies (Davis & Widseth, 1978; Strupp & Bloxom, 1975) the Depression scale (2) was a statistically significant predictor of GGPA. This result was especially noteworthy because results were replicated using the MMPI-2 rather than the original MMPI. Specifically, results from this study indicated that a one standard deviation decrease in score on the Depression scale (2) would yield a .24 increase in GGPA. According to Graham (2006), people who score high on the Depression scale "tend to feel hopeless and to be pessimistic about the future in general and more specifically about the likelihood of overcoming their problems and making better adjustment" (p. 69). Given that professional counselors aim to instill hope and assist people make better adjustment, the negative relationship seems logical. Findings also imply that students with greater levels of depression may have more difficulty focusing on academic tasks than those with fewer symptoms of depression.

As predicted in H2, the MMPI-2 showed incremental validity, over UGPA and PWR, in the prediction of internship evaluation ratings, but not at the .05 alpha level set for the study. The total model accounted for 15.6% of the outcome variance and of that, 14.5% was explained by the MMPI-2 predictor variables. Given that the p values were not significant for any of the models at the .05 alpha level set for this study, a statistical equation could not be built. However, the final model with all three blocks was significant at the .10 alpha level, which indicated that there is a one in ten chance that a researcher will reject the null hypothesis when the null hypothesis is in fact true.

Specifically, Scale *F* (p = .023) and Scale 2 (p = .002) on the MMPI-2 each significantly contributed the prediction of IER at the .05 alpha level. Scale F is called the Infrequency Scale and high scores are typically associated with individuals with severe pathology or with a desire to appear more psychologically disturbed than they are in actuality (Graham, 2006). Given that that Scale F and IER were positively associated, it meant that as the Infrequency scores increase so do the IERs. It is notable; however, that the mean *T*- score (*T*=43; *SD*=6) for the participants was well within the average range. According to Graham (2006), scores in this range for nonclinical populations indicate that the respondent answered in an acceptable manner and may be endorsing some deviant beliefs. "Sometimes scores at this level indicate persons with deviant social, political, or religious convictions. Persons with scores at the upper end of this range may be accurately reporting psychological problems"(Graham, 2006, p. 29). Thus, the relative elevations may actually make the student better able to empathize with a client's emotional struggles.

As previously mentioned, Scale 2 is the Depression scale and given that the findings are similar to those found when testing H1, similar conclusions can be drawn. Participants who endorsed more symptoms associated with depression at the beginning of the PCMP tended to have lower IERs at the end of the PCMP. These findings suggest that due to increased symptoms of depression, participants may have had difficulty in the role of a counselor or that they may have had more difficulty focusing on academic tasks than those with fewer symptoms.

In a similar study using the MMPI and internship evaluations, Daehnert and Carter (1987) found very different results. Correlations between MMPI-2 scale scores

and internship evaluations ranged from -.21 to .56 and in this study the strongest correlation was -.21. Given the sample populations were both graduate psychology students, it seems likely that different results were achieved because this study used the mean score from internship evaluations rather than scores from distinct categories. As Oswald et al. (2004) stated, "The number of dimensions should not be so many that the information is unwieldy, yet not so few that the domain of college performance is oversimplified and not appropriately represented" (p.187). It appears possible that using the mean score oversimplified the criterion and that stronger correlations may have become apparent if the multidimensional evaluation categories had been utilized.

Given that the MMPI-2 performed as predicted but not at the .05 alpha level set for the study, it is important to discuss the implications of the H2 results in terms of practical verses statistical significance. According to Kirk (1996), "Statistical significance is concerned with whether a research result is due to chance or sampling variability; practical significance is concerned with whether the result is useful in the real world"(p.746). He argued that the myopic focus on null hypothesis significance testing is problematic in professional psychology research. He stated, "Unfortunately, all too often the primary focus of research is on rejecting a null hypothesis and obtaining a small *p* value. The focus should be on what the data tell us about the phenomenon under investigation" (Kirk, 2001, p.213). Therefore, the practical significance of the results for H2 are that the MMPI-2 can provide incremental validity and predict performance on IER in a PCMP beyond what can be accounted for by traditional admissions criteria. Additionally, it lends empirical support to previous findings that the predictive validity of personality increases at higher levels of education where the pool of applicants is more

homogenous and the criterion of success involves application of knowledge to applied practice. Finally, it demonstrated that in addition to personality instruments based on the Five Factor Model (FFM), the MMPI-2 can also provide incremental validity in the prediction of academic outcomes.

H3 predicted that a stronger statistical model and more outcome variance could be explained when the criterion of success was more in line with the training objectives at the PCMP. This hypothesis was inspired by the well documented issue in predictor – criterion and incremental validity research.

> Within the complex and competitive admissions process, colleges and universities seek out the best students possible for their institutions, in which "best" can be defined in many ways….However, critics argue there is substantial room for improvement with respect to the validity and practical utility of current selection tools (Oswald et al., 2004, p.187).

In both analyses, the MMPI-2 had the most independent and incremental effects on the outcome. However, the most variance was explained when GGPA was the criterion and when IER was the criterion none of the models were statistically significant at the .05 alpha level set for the study. Therefore, H3 was rejected.

Implications of the Study and for Professional Counseling Masters Programs

The results of this study contribute to the literature in several significant ways. First, it added to the sparse research on predictive validity at the graduate level and the expressed need to identify other criterion of graduate education success as well as other predictor variables of success in graduate school. Second, existing research on personality traits and academic success had used instruments based on the Five Factor Model (FFM) and were conducted for research purposes only. Therefore this study

diversified that line of inquiry and evidenced that like instruments based on the FFM (Chamorro-PreMuzic & Furnham, 2003; Trapmann, Hell, Hirn & Schuler, 2007), the MMPI-2 has independent and incremental effects on academic outcomes, after traditional predictors of those outcomes are controlled. In fact, the MMPI-2 demonstrated its utility for admission decisions given that the MMPI-2 was still able to significantly contribute to the prediction model irrespective of the inherent range restriction of scores that resulted from using a sample already prescreened with the predictor variables. Likewise, this study confirmed that MMPI-2 scores on the Depression (2) and Hypomania (9) scales are indicators of future performance (2-3 years after the MMPI was completed). Accordingly, this study has provided empirical evidence to support the utility of including valid, reliable personality instruments in PCMP admissions procedures.

It is beyond the scope of the results to ascertain with certainty how well PWR would have predicted performance in a PCMP without the restriction of range of scores. However, it is quite likely that had the PWR ratings not been used to make admissions decisions (i.e., majority of accepted applicants had high scores), these ratings would have accounted for more of the variance than found in this study. Accordingly, results should be interpreted with the caveat that the results show a lack of predictive validity for PWR only within a restricted range of scores. Additionally, it highlights once again, that even with the inherent range restriction that results when using participants who have been previously screened and admitted into a competitive graduate programs, the MMPI-2 performed well and demonstrated utility for admissions procedures. This could be of interest to applied researchers and admissions committees given that subjective assessments of personality characteristics (e.g., letters of recommendation, interview, and

personal statements) are the most common way admissions committees screen (Johnson & Cambell, 2004; Norcross, Kohout& Wicherski, 2005).

There are several additional implications from this study for PCMPs. First, empirical support for the inclusion of an objective personality assessment as part of standard admissions in a PCMP was demonstrated. In this study, the MMPI-2 was the instrument used in the predictive model; however, instruments based on the Five Factor Model have also been shown to provide incremental validity of personality traits in the prediction of academic outcomes (Poropat, 2009; Trapmann et al., 2007). Given that the MMPI-2 has 567 questions and takes between 60-90 minutes to complete (Weiner & Greene, 2008) it may be worthwhile for PCMP to test whether personality instruments based on the FFM with fewer questions, such as the NEO-PI-R (240 questions) or the NEO-FFI (60 items) (Costa & McCrae, 1991, 1992) can predict a commensurate amount of variance in performance as the MMPI-2 did in this study. This may improve the utility or practicality of incorporating the findings of this study into standard admissions procedures at other PCMPs.

Finally, PCMPs should consider the practical significance of the finding that participants who scored higher on the Depression Scale (2) and the Hypomania Scale (9) had lower performance outcome scores. It is important to underscore that relatively speaking, even those participants who scored higher on Scales 2 and 9, persevered well in the program as evidenced by graduating. Therefore, it is suggested that applicants with higher scores on Scales 2 and 9 be identified in order to direct them towards resources that can assist improve their mental health rather than as a method to screen out the applicant from admissions. This seems especially important given that as individuals

progress in graduate psychology programs, the overall number of distressed students increase and the level of severity of their distress increases as well (Kuyken, Peters, Power & Lavender, 2003).

Limitations and Directions for Future Research

There were several limitations to this study. Accordingly, these will be reviewed and possible remedies will be offered. First, this study was retrospective; therefore, the analysis was bound by the data that had been collected as part of admissions at the PCMP which could affect the ability to generalize results to PCMPs with different admissions criteria. Accordingly, future research could replicate this study with different cognitive as well as non cognitive predictors and the MMPI-2. For example, a multivariate analysis using GRE as the cognitive variable in the first block, letters of recommendation as the subjective assessment of personality traits in the second block and the MMPI-2 in the third block.

Second, the participants had been screened and admitted to the PCMP which likely affected the range of scores. Similarly these participants were seeking admissions to a competitive MA program and may have attempted to present themselves in a particularly favorable manner which may not have accurately reflected actual abilities or traits. Given that the participants would generally be considered high achieving, it is recommended that the study be replicated and that range restriction be statistically accounted for. However this possibility is somewhat mitigated by the fact that if an applicant had a MMPI-2 profile that was considered to be invalid, the applicant was required to retake the instrument provided that demographic variables (e.g., applicant spoke English as a second language) were not a possible factor.

Third, the participants all attended a PCMP at a medium university in the Rocky Mountain region. Given that the findings may not generalize to other countries or regions of the United States, it is recommended that this study be replicated at PCMPs in other regions.

Fourth, there were delimitations in this study, which primarily related to the design. The decision to delimit the MMPI-2 variables to the ten basic profiles and three of the nine validity scales was guided by parameters of the multivariate analysis. Based on the significant findings of this study, future researchers could use the MMPI-2 subscales, such as the Harris-Lingoes, in an analysis to better understand how the dimensions of Depression and Hypomania relate to performance in a PCMP. Similarly, it would be interesting to go beyond the 13 scales used in this study and examine configural aspects of the MMPI-2 scale scores. For example, an examination of code types or of cases where Scale1 and Scale 3 are both ten T-score points higher than Scale 2 (i.e., suggestive of somatization) (Graham, 2006) might provide a better understanding as to how personality and performance in a PCMP are related.

Fifth, it is important to address how using grade point average as variables likely affected the results. In terms of graduate grade point average as the criterion, the use of a sample of graduate students from a single academic discipline, with a standardized curriculum with which to compare grade point averages was more reliable and consistent than would be the case if different academic disciplines or programs from other universities were part of the sample. Conversely, the use of undergraduate grade point average did not have the same benefits. As Oswald et al. (2004) explained, “GPA as a composite measure is not standardized and may represent the outcome of some very

different student behaviors, as reflected in different types of courses taught by different instructors"(p. 187). Therefore, this lack of standardization is a caveat because it is difficult to assess whether there are true differences in scores. Additionally, it is doubtful that the UGPA is a reflection of purely cognitive abilities. It would be beneficial if this study was replicated either substituting the cognitive variable with GRE scores or adding a block for the GRE scores to better understand if some of the variance explained by undergraduate GPA could in actuality be more related to personality. In other words, one's GPA reflects more than just intellect whereas the GRE seems to be more of a quantitative and objective assessment of cognitive abilities.

Lastly, the findings of H2 could be interpreted to suggest that the MMPI-2 measured characteristics that are distinct from those measured by internship evaluations. However, given the findings by Daehert and Carter (1987) it is also possible that the statistical analysis did not allow for a more comprehensive examination of the relationship between personality factors and internship skills. In other words, the mean score of the internship evaluation ratings may have been too narrow and an analysis that could examine subcategories of the IER as multiple dependent variables might provide a more accurate depiction of the correlations. For instance, what scales and T – score elevations are associated with being open to supervision/feedback or with ethical practice? This could be achieved by computing Pearson correlation coefficients between each of the MMPI-2 scales and the sub category mean scores on the internship evaluations.

Summary

The purpose of this study was to investigate whether there is empirical support to include an objective measure of personality characteristics as part of standard admissions procedures in a professional counseling masters program (PCMP). As hypothesized, the MMPI-2, an objective personality instrument, had incremental and independent effects on two academic outcomes; namely, graduate grade point average (GGPA) and internship evaluation ratings (IER). In fact, both analyses evidenced that after controlling for traditional admissions criteria, undergraduate grade point average (UGPA) and Pre-Admissions Workshop ratings (PWR), the MMPI-2 accounted for the most variance in both outcome variables. These results imply that personality characteristics are veritably relevant to academic performance and advance prior research indicating that objective measures of personality are predictive of academic outcomes. As Poropat (2009), who conducted the most comprehensive meta-analytic investigation of personality-academic performance relationships concluded, "Personality should take a more prominent place in future theories of academic performance and not merely as an adjunct to intelligence" (p. 333). The third hypothesis sought to identify criterion other than GGPA as a measure of success and tested whether a stronger statistical model could be built when using a criterion more in line with the training objectives of a PCMP. This was rejected because more variance was explained in the model using GGPA and an equation could not be built when using IER. However, using IER as a criterion of success in a PCMP did demonstrate potential to overcome the limitations of using GPA and future research should focus on a way to capture the different dimensions of internship performance when analyzing it as a criterion. Overall, the practical significance of the

present study is that it moves the field of professional psychology closer towards developing improved admissions criteria for graduate school admissions and underscores the importance of including personality characteristics when building models to predict performance outcomes.

REFERENCES

American Psychological Association Board of Educational Affairs. (2007). *Assessment of competency benchmarks work group: A developmental model for the defining and measuring competence in professional psychology.*

Banks, Mooney, Mucowski, & Williams, (1984). Progress in the evaluation and prediction of successful candidates for religious careers. *Counseling and Values, 28,* 82-91. Retrieved from http://www.counseling.org/Publications/Journals.aspx

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, achievement motivation and academic success of psychology students in higher education. Personality and Individual Differences, *29*, 1057–1068. doi:10.1016/S0191-8869(99)00253-6

Butcher, J.N. (1979). Use of the MMPI in personnel selection. In Butcher, J.N. (Ed.), *New developments in the use of the MMPI*. Minneapolis: University of Minnesota Press.

Butcher, J.N. (1985). Personality assessment in industry. In Bernardin, H.J., & Bownas, D. (Eds.), *Personality assessment in organizations.* New York: Praeger.

Butcher, J.N. (1991). Screening for psychopathology: Industrial applications of the Minnesota Multiphasic Personality Inventory (MMPI-2). Chapter in J. Jones, B.D. Steffey, & D. Bray (Eds.). *Applying psychology in business: The manager's handbook.* Boston: Lexington Books.

Butcher, J. N., Ones, D.S. & Cullen, M. (2006). Personnel screening with the MMPI-2. In J.N. Butcher (Ed.). *MMPI-2: The practioner's handbook*. (pp. 381-406) Washington, D.C.: American Psychological Association. doi:10.1037/11287-014

Butcher, J.N. & Rouse, S.V. (1996). Personality: Individual differences and clinical assessment. *Annual Review of Psychology*, *47*, 87-111. doi:10.1146/annurev .psych.47.1.87

Callan, R.J. (1997). An attempt to use the MMPI as a predictor of failure in military training. *Military Psychology, 121,* 533-557. Retrieved from http://0-www .informaworld.com.source.unco.edu/smpp/title~content=t775653681~db=all

Celmer, V & Winer, J. L. (1990). Female aspirants to the Roman Catholic priesthood. *Journal of Counseling and Development, 69,* 178-183. Retrieved from http://www.counseling.org/Publications/Journals.aspx

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Psychology,*37*,* 319–338. doi:10.1016/S0092-6566(02)00578-0

Chamorro-Premuzic, T., & Furnham, A. (2006). Intellectual competence and the intelligent personality: A third way in differential psychology. *Review of General Psychology*, *10*(3), 251-267. doi: 10.1037/1089-2680.10.3.251

Chibnall, J.T., & Detrick P. (2003). The Neo PI-R Inwald Personality Inventory and MMPI-2 In The Prediction Of Police Academy Performance: A Case for Incremental Validity. *American Journal of Criminal Justice, 27,* 233-248. doi:10.1007/BF02885696

Compton, K. (2009). *Using the MMPI-2 as a predictor of academic success* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses database. (UMI No. 3352025)

Costa, P.T., & McCrae, R.R. (1991). *The NEO Five Factor Inventory*. Odessa, FL: Psychological Assessment Resources.

Costa, P.T., & McCrae, R.R. (1992). *Revised NEO Personality Inventory and NEO Five Factor Inventory Professional manual.* Odessa; FL: Psychological Assessment Resources.

Daehnert, C., & Carter, J.D. (1987). The prediction of success in a clinical psychology graduate program. *Educational and Psychological Measurement*, *47*, 1113-1125. doi:10.1177/0013164487474029

Detrick, P., Chibnall, J.T., & Rosso, M. (2001). Minnesota Multiphasic Personality Inventory—2 in police officer selection: Normative data and relation to the Inwald Personality Inventory. *Professional Psychology: Research and Practice, 32*(5)*,* 484-490. doi:10.1037//0735-7028.32.5.484

Di Fabio, A. & Palazzeschi, L.(2009). An i-depth look at scholastic success: Fluid intelligence, personality traits or emotional intelligence? *Personality and Individual Differences*, *46*, 581-585. doi:10.1016/j.paid.2007.06.025

Dunnette, M.D., Bonas, D.A., & Bosshardt, M.J. (1981). *Electric Power Plant Study: Prediction of inappropriate, unreliable, or aberrant job behavior in nuclear power plant settings.* Minneapolis: Personnel Decisions Research Institute.

Epstein, R.M. & Hundert, E.M. (2002) Defining and assessing professional competence *JAMA*.287:226-235. doi:10.1001/jama.287.2.226

Farsides, T. & Woodfield, R. (2003). Individual differences and undergraduate academic success: The roles of personality, intelligence, and application. *Personality and Individual differences, 34*(7), 1225-1243. doi:10.1016/S0191-8869(02)00111-3

Furnham, A., Monsen, J. & Ahmetoglu, G. (2009). Typical intellectual engagement, big five personality traits, approaches to learning and cognitive ability predictors of academic performance. *British Journal of Educational Psychology*, *79*, 769-782. doi:10.1348/978185409X412147

Graham, J.R. (1988). *Establishing validity of the revised form of the MMPI.* Symposium presentation at the 96th Annual Convention of the American Psychological Association, Atlanta, GA.

Graham, J.R. (2006) *MMPI-2 Assessing personality and psychopathology*. New York: Oxford University Press.

Hargrave, G.E., Hiatt, D., Gaffney, T.W. (1988). F+4+9+Cn: An MMPI measure of aggression in law enforcement officers and applicants. *Journal of Police Science & Administration,16*, 268-273.

Haynes, S.N. & Lench, H.C. (2003). Incremental validity of new clinical assessment measures. *Psychological Assessment*, *15*, 4, 456-466. doi:10.1037/1040 -3590.15.4.456

Hiatt, D., & Hargrave, G. E. (1988). MMPI profiles of problem peace officers. *Journal of Personality Assessment. 52* (4), 722-731. Retrieved from http://0-www .informaworld.com.source.unco.edu/smpp/title~content=t775653663~db=all

Hunsley, J. (2003). Introduction to the special section on incremental validity and utility in clinical assessment. *Psychological Assessment, 15*(4), 443-445. doi:10.1037 /1040-3590.15.4.443

Hunsley, J. & Meyer, G.J. (2003) The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment*, *15* (4), 446-455. doi: 10.1037/1040-3590.15.4.446

Jansen, D.G., Bonk, E.C., & Garvey, F.J. (1973). MMPI characteristics of clergymen in counseling training and their relationship to supervisors' and peers' ratings of counseling effectiveness. *Psychological Reports, 33,* 695-698. Retrieved from http://www.ammonsscientific.com

Johnson, W. & Campbell, B. (2004). Character and fitness requirements for professional psychologists: Training directors' perspectives. *Professional Psychology: Research & Practice*, 35, 405-411. doi:10.1037/0735-7028.35.4.405

Kelly, W.L. (1974). Psychological prediction of leadership in nursing. *Nursing Research, 23,* 83- 42. doi:10.1097/00006199-197401000-00009

King, A. R., & Bailly, M. D. (2002). MMPI-2 predictors of academic performance. *Journal of Psychoeducational Assessment, 20*(3), 258-267. doi:10.1177 /073428290202000303

Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, *56* (5), 746-759. doi:10.1177 /0013164496056005002

Kirk, R. E. (2001). Promoting good statistical practices: Some suggestions. *Educational and Psychological Measurement, 61* (2), 213-218. doi:10.1177 /00131640121971185

Kodman, F. (1984). Some personality traits of superior university students. *Social Behavior and Personality, 12* (2), 135-138. doi:10.2224/sbp.1984.12.2.135

Kuyken, W., Peters, E., Power, M. J., & Lavender, T. (2003). Trainee clinical psychologists' adaptation and professional functioning: A longitudinal study. *Clinical Psychology &Psychotherapy, 10*, 41–54. doi:10.1002/cpp.350

Lachar, D., (1974). Prediction of early US Air Force freshman cadet adaptation with the MMPI. *Journal of Counseling Psychology, 21,* 404-408. doi:10.1037/h0037112 .4.291

Leverett-Main, S. (2004). Program directors'perceptions of admission screening measures and indicators of student success. *Counselor Education & Supervision*, *43*, 207-219. Retrieved from ttp://www.counseling.org/Publications/Journals.aspx

Lievens, F., Dilchert, S. & Ones, D.S. (2009). Personality scale validities increase throughout medical school. *Journal of Applied Psychology*, *94* (6), 1514-1535. doi:10.1037/a0016137

Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology, 89* (2), 187–207. doi: 10.1037/0021-9010.89.2.187

Pallant, J. (2007). SPSS Survival Manual. Open University Press: New York, NY.

Poropot, A.E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135* (2), 322-338. doi:10.1037 /a0014996

Rouse, S. V. & Butcher, J. N. (1995). *Annotated bibliography on the use of the MMPI/MMPI-2 in personnel and educational selection.* Minneapolis, MN: University of Minnesota Press.

Shusman, E.J., Inwald, R.E., & Kantz, H.F. (1987). A cross validation study of police recruit performance as predicted by the IPI and MMPI. *Journal of Police Science and Administration, 15*, 162-169.

Sternberg, R.J. & Williams, W.M. (1997). Does the graduate record examination predict meaningful success in the graduate training of psychology? A case study. *American Psychologist, 52*, 630-641. doi:10.1037//0003-066X.52.6.630

Strupp, H.H.& Bloxom, A.L. (1975). An approach to defining a patient population in pscyhotherapy research. *Journal of Counseling Psychology*, *3*, 231-237. doi:10.1037/h0076881

Trapmann, S., Hell, B., Hirn, J.W., & Schuler, H. (2007). Meta-Analysis of the relationship between the big five and academic success at university. *Journal of Psychology*, *215*, 132-151. doi.10.1027/0044-3409.215.2.132

Weiner, I.B., & Greene, R.L. (2008). Handbook of personality assessment. New York: Wiley.

Weiss, W.U., Serafino, G., Serafino, A., Willson, W. & Knoll, S. (1998). Use of the MMPI-2 to predict the employment continuation and performance ratings of

recently hired police officers. *Journal of Police and Criminal Psychology*, *13*(1), 40-44. doi:10.1007/BF02819065

Weiss, W.U., Serafino, G., Serafino, A., Willson, W., Sarsany, J. & Felton, J. (1999). Use of the MMPI-2 and the Inwald Personality Inventory to identify the personality characteristics of dropouts from a state police academy. *Journal of Police and Criminal Psychology*, *14* (1), 38-42. doi:10.1007/BF02813861

Wheeler, S. (2000). What makes a good counselor? An analysis of ways in which counselor trainers construe good and bad counseling trainees. *Counseling Psychology Quarterly,13*(1), 65-83. doi:10.1080/09515070050011079

APPENDIX B

PRE-ADMISSIONS WORKSHOP RATING FORM

Evaluative Criteria for MA Workshops

Applicant's Name: _____

Reviewer's Name: _____

Please use the following scale:

NA: Behavior not observed
1- Inadequate or poor
2- Fair
3- Average
4- Good

1.	Tolerance of individual differences (no readily apparent biases)	1	2	3	4
	NA				
2.	Able to intelligently and thoughtfully Discuss the concerns expressed by the 'client' in the role play.	1	2	3	4
	NA				
3.	Interacts appropriately with other group members.	1	2	3	4
	NA				
4.	Maturity	1	2	3	4
	NA				
5.	Please rate how successful this individual Would be in a practicum class (i.e., 612)	1	2	3	4
	NA				

6. Strengths_____

_____

7. Weaknesses or Concerns_____

_____

8. Please provide an <u>overall</u> rating of This individual	1	2	3	4

Additional Comments (if applicable):

APPENDIX C

INTERNSHIP EVALUATION RATING FORM

ATTACHMENT D

INTERNSHIP STUDENT EVALUATION
APCE 692 Clinical Counseling
School of Applied Psychology and Counselor Education

Name of Student:_____Date:_____

Name of Supervisor:_____Site:_____

Address:_____

_____

Please describe your style of supervision with this student.

_____Observed student directly
_____Listened to or watched tapes of student counseling (minimum of 3)
Dates: _____ _____ _____ _____
_____Read session notes
_____Discussed cases with student
_____Other (please describe)

What number of overall Internship hours did the student spend with:

_____ Direct client activities (counseling)
_____ Indirect client activities (i.e. case conferences, staff meetings, administrative duties, etc.)

In order to facilitate the development of the student's skills, please rate student's performance according to the following scale:

Low				High	
1	2	3	4	5	N/A (if behavior was not observed)

Please write comments, when necessary, in the space provided under each question. Your feedback is greatly appreciated.

1) OPENING / RAPPORT

* Did counselor establish good rapport with client? 1 2 3 4 5 N/A

2) INTERACTION / INTERVIEW SKILLS

* Was counselor in control of direction of interview or did clients go off on meaningless tangents?	1	2	3	4	5	N/A
* Was counselor accepting and encouraging of client's emotions, feelings, and expressed thoughts?	1	2	3	4	5	N/A

3) COUNSELOR RESPONSES

* Were counselor's responses appropriate in view of what counselee was expressing?	1	2	3	4	5	N/A
* Did counselor reflect and react to feelings?	1	2	3	4	5	N/A
* Did the counselor's values remain objective when working with the client?	1	2	3	4	5	N/A
* Were interventions used appropriately?	1	2	3	4	5	N/A

4) COUNSELING RELATIONSHIP

* Was relationship conducive to productive counseling?	1	2	3	4	5	N/A
* Did counselor talk at appropriate language level with clients?	1	2	3	4	5	N/A
* Did the counselor use language, tone of voice, and other behavior to convey an interest in the client?	1	2	3	4	5	N/A
* Did counselor communicate his/her interests, feelings and experiences to the client when appropriate?	1	2	3	4	5	N/A

5) CLIENT CONCEPTUALIZATION

* Did counselor understand/conceptualize client's problem in its full perspective (i.e. systems)?	1	2	3	4	5	N/A
* Can counselor report client behavior accurately and support observations with specific behavioral observations?	1	2	3	4	5	N/A
* Did interventions reflect a clear understanding of the client's problem?	1	2	3	4	5	N/A
* Was counselor able to demonstrate knowledge of principles and processes of theoretical framework underlying mode of treatment used?	1	2	3	4	5	N/A
* Were treatment goals and plans reflective of good case conceptualization?	1	2	3	4	5	N/A

6) TERMINATION

* Was termination initiated properly (was it a smooth transition from the counseling process)?	1	2	3	4	5	N/A
* Was follow up or termination discussed?	1	2	3	4	5	N/A

7) CASE CONCEPTUALIZATION / SUPERVISION

* Was counselor able to observe and discuss case objectively and insightfully with supervisor?	1	2	3	4	5	N/A
* Was counselor receptive to supervisor feedback?	1	2	3	4	5	N/A
* Was feedback reflected in future counseling sessions?	1	2	3	4	5	N/A
* Was counselor able to observe/understand his/her own personal influence on the counseling relationship?	1	2	3	4	5	N/A

APPENDIX D

UNIVERSITY OF NORTHERN COLORADO INSTITUTIONAL REVIEW BOARD APPROVAL

UNIVERSITY of NORTHERN COLORADO
Institutional Review Board (IRB)

April 7, 2011

TO: Maria Lahman
Applied Statistics and Research Methods

FROM: The Office of Sponsored Programs

RE: Exempt Review of *Predictors of Performance in a Professional Counseling Masters Program,* submitted by Mya Bethune (Research Advisor: Brian Johnson)

The above proposal is being submitted to you for exemption review. When approved, return the proposal to Sherry May in the Office of Sponsored Programs.

I recommend approval.

M.J.S.K 4-15-11
Signature of Co-Chair Date

The above referenced prospectus has been reviewed for compliance with HHS guidelines for ethical principles in human subjects research. The decision of the Institutional Review Board is that the project is exempt from further review.

IT IS THE ADVISOR'S RESPONSIBILITY TO NOTIFY THE STUDENT OF THIS STATUS.

Comments:

25 Kepner Hall ~ Campus Box #143
Greeley, Colorado 80639
Ph: 970.351.1907 ~ Fax: 970.351.1934